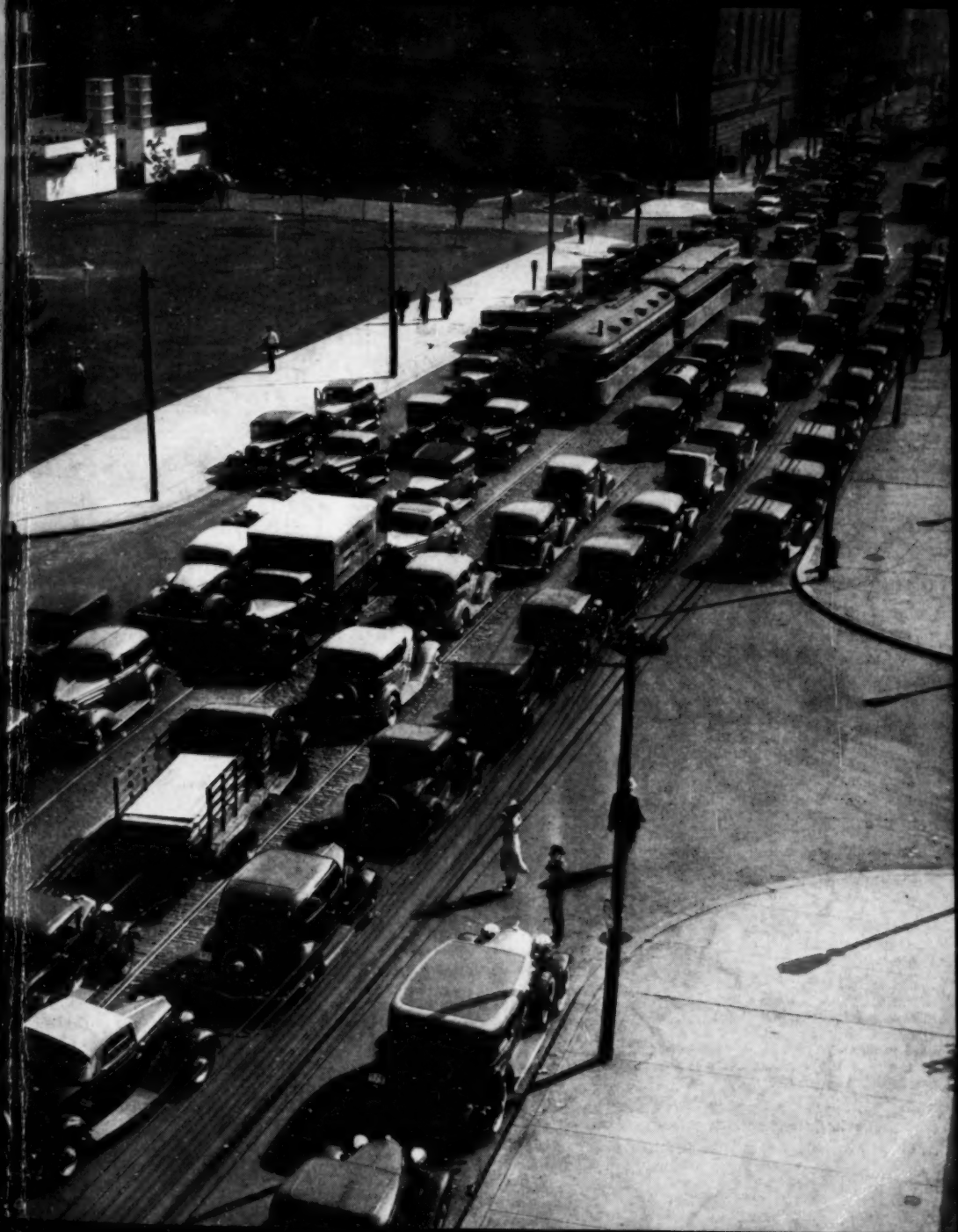
JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED PHYSICS

Vol. 23, No. 2 FEBRUARY, 1952

In This Issue

		Page
Operations Research, What Is It?	PHILIP M. MORSE	165
Contributed Original Research		
On the Accumulation of Errors in the Numerical Solution of Differential Equations	A. PAPOULIS	173
Thermal Conductivity of Metals at High Temperature	C. L. HOGAN AND R. B. SAWYER	177
Cathodic Sputtering for Micro-Diffusion Studies	THOMAS F. FISHER AND C. E. WEBER	181
Unsteady-State Separation Performance of Cascades. I	ELLIOTT W. MONTROLL AND GORDON F. NEWELL	184
An Iterative Method for the Solution of Eigenvalue Problems	V. J. BERRY AND C. R. DE PRIMA	195
Phenomena Associated with the Flight of Ultra-Speed Pellets. Part II. Spectral Character of Luminosity	W. C. WHITE, JOHN S. RINEHART, AND WILLIAM A. ALLEN	198
Reversion Method for Solving Nonlinear Differential Equations	LOUIS A. PIPES	202
Space Charge and Transit Time Considerations in Planar Diodes for Relativistic Velocities	HENRY F. IVEY	208
Structure of Tetragonal Tungsten Trioxide	W. L. KEHL, R. G. HAY, AND D. WAHL	212
Space-Charge-Wave Propagation in a Cylindrical Electron Beam of Finite Lateral Extension	PHILIP PARZEN	215
Dielectric Constant of Water Vapor in the Microwave Region	GEORGE BIRNBAUM AND S. K. CHATTERJEE	220
Generalized Ideal Filters	LOTFI A. ZADEH AND KENNETH S. MILLER	223
Effect of Gas Pressure in the Electric Bulb on the Polarity of the Rectified Current	SHANG-YI CH'EN AND WEI-CHENG LI	228
Considerations Regarding the Distribution of Radiation Damage in Matter Bombarded by Light Ions and a Method of Making it Uniform	PHILIP SCHWED AND GERHART GROETZINGER	234
Studies in Newtonian Flow. III. The Dependence of the Viscosity of Liquids on Molecular Weight and Free Space (in Homologous Series)	ARTHUR K. DOOLITTLE	236
Space-Charge-Limited Currents between Inclined Plane Electrodes	HENRY F. IVEY	240
Diffraction Measurements at 1.25 Centimeters	RALPH D. KODIS	249
Emission of Radiation from Diatomic Gases. III. Numerical Emissivity Calculations for Carbon Monoxide for Low Optical Densities at 300°K and Atmospheric Pressure	S. S. PENNER, M. H. OSTRANDER, AND H. S. TSIEN	256
Secondary Electron Emission from Metals under Positive Ion Bombardment in High Extractive Fields	E. W. WEBSTER, R. J. VAN DE GRAAFF, AND J. G. TRUMP	264
Measurement of Ultrasonic Attenuation in Solids by the Pulse Technique and Some Results in Steel	R. L. RODERICK AND JOHN TRUELL	267
Electric Field in Diodes and the Transit Time of Electrons as Functions of Current	PAUL L. COPELAND AND DELBERT N. EGGENBERGER	280
Letters to the Editor:		
Orientation of Crystallites in Stretched Polyethylene	SAMUEL KRIMM	287
Reply to Samuel Krimm's Note	ALEXANDER BROWN	287
Distribution of Activation Energies for Viscoelastic Deformation	W. JAMES LYONS	287
Viscous Energy Dissipated During the Atomization of a Liquid	G. W. MONK	288
Note on "Electron Microradiography of Electrodeposited Metals"	SHIGETO YAMAGUCHI	288
Radiation or Diffraction Patterns Close to Receiving Antennas	L. S. PALMER	289
Concerning the Letter by Mr. P. E. Ohmart on "A Method of Producing Electric Current from Radioactivity"	W. H. WALDO	290
Books Reviewed		291
Other Books Received		296
Index to Advertisers		xxviii

BOARD OF EDITORS

ELMER HUTCHISSON, *Editor*

EARLE C. GREGG, *Assistant Editor*

Associate Editors

HOWARD H. AIKEN
RAY E. BOLZ
LYLE D. BORST
ROBERT D. HEIDENREICH
JAMES HILLIER
LEONARD B. LOEB
W. JAMES LYONS
ERIC REISSNER
ROMAN SMOLUCHOWSKI

Publication of the

American Institute of Physics

OFFICERS

GEORGE R. HARRISON, *Chairman*
GEORGE B. PEGRAM, *Treasurer*

GOVERNING BOARD

J. W. BEAMS
C. PAUL BONER
WALLACE R. BRODE
W. F. FAIR, JR.
S. A. GOUDSMIT
G. P. HARNWELL
GEORGE R. HARRISON
F. V. HUNT
RUDOLPH KINGSLAKE
PAUL KIRKPATRICK
F. W. LOOMIS
J. R. OPPENHEIMER
GEORGE B. PEGRAM
I. I. RABI
DUANE ROLLER
HAROLD K. SCHILLING
FREDERICK SEITZ
FRANCIS G. SLACK
JOHN C. STEINBERG

Consultant on Printing

MELVIN LOOS

ADMINISTRATIVE STAFF

HENRY A. BARTON, *Director*
WALLACE WATERFALL, *Executive Secretary*
THEODORE VORBURGER, *Advertising Manager*
RUTH F. BRYANS, *Publication Manager*
BETTY H. GOODFRIEND, *Assistant Secretary*
EILEEN B. NEUBERGER, *Circulation Manager*
KATHRYN SETZE *Assistant Treasurer*

The JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. The JOURNAL OF APPLIED PHYSICS publishes editorials and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. Unless otherwise stated, submission of a manuscript is a representation that it has been neither copyrighted, published, nor submitted for publication elsewhere. The authors' institutions are requested to pay a **publication charge** of $8 per page which, if honored, entitles them to 100 reprints without further charge. Instructions will be sent with galley proofs.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members of the American Institute of Physics and Affiliated Societies	$10.00	$11.00
To all others	12.00	13.00

Back Number Prices

Yearly back number rate when complete year is available: $10.00. Single copies: $1.00 each.

Changes of address should be sent to the American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics. **Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Copyright 1952, by the American Institute of Physics.

Journal
of
Applied Physics

Volume 23, Number 2 February, 1952

Operations Research, What Is It?*

PHILIP M. MORSE

Massachusetts Institute of Technology, Cambridge, Massachusetts

THE question in the title of this talk is a real one, not a rhetorical one. Here is an activity[1] which helped save lives and helped win battles during the war, which is now included as a matter of course in all the staffs of all the major Atlantic pact military forces, which is being used with success in an increasing number of industrial staffs; yet discussions as to what it is seem to come to nothing more definite than fist fights.

Personally, I would prefer to forget about definitions and get on with the work. After all, who cares what it's called, what it is, or what it is not, as long as it's useful and is used. But I am afraid we can't quite let it go at that. For one thing, operations research is growing faster than there are experienced men available. If the work is to develop in a healthy way, more men must be trained. In order to do this we will have to decide how best to do the training. Is on-the-job training the only sort of training which counts, and must we continually bleed our present operations research teams to start new ones? Or can we at least partially prepare men in college for this work? At least three large universities are now proposing to find the answer to this question. To do it they must decide what training for operations research should include.

In addition, many industrial firms are just learning about operations research. In order to determine more clearly the limitations as well as the possibilities of the technique, some of these firms should try it out. They certainly won't buy a pig in the dark, so to speak; they must know what it is in order to try it out effectively.

A DEFINITION IS DIFFICULT

I suppose it was nearly as difficult to define any branch of science when it was started. For that matter, can you define what physics is right now—or chemistry? Try it. Chemists study the reactions between substances? So does the housewife when she cooks. Physicists study the forces between bodies? So do wrestlers. Operations research can't be defined entirely in terms of subject matter either. The operations research worker studies the behavior of organizations, the operations of men and equipment in combination, but so does any competent administrator, so do efficiency experts. Operations research involves a study of problems of interest to management, but certainly these are the problems which interest any management expert. Evidently such a definition is too broad; it doesn't differentiate.

Part of the definition must be in the way the problems are looked at. The housewife is not a chemist unless she becomes more interested in the baking of the cake than in the eating, unless she tries to find out just why the cake comes out the way it does and unless she carries out her investigations in a particular way. For instance, if she tries to find out the relationship between baking temperature and the texture of the cake, she will turn into a physicist rather than a chemist! The physicist, or the chemist, employs a characteristic method of attack on a problem, uses certain techniques which by now are known well enough to be capable of being taught in regular courses. The physical scientist uses the quantitative language of mathematics, employs the well-known but difficult-to-describe procedures of experimentation and theory-making. He looks at the phenomenon he is studying in a certain impersonal way being more interested in *how* than in *whither*, more interested in *why* than in *for what use*. Many centuries of experience have taught him that this impersonal viewpoint, this dual employment of theory and experiment, will usually procure for him results of value in his science and that too great a preoccupation with questions of the worth of the result, or the immediacy of the need,

* Presented at a conference on Applications of Operations Research in Industry at Case Institute of Technology, Cleveland, Ohio, November 8–10, 1951.

will actually hinder his progress. This is not to say that questions of worth or immediacy are trivial, or to say that the results of science are the only worthwhile results to obtain. Far from it. What I do wish to say is that the results of scientific activity have not been without value and that to obtain these results one must use the techniques and viewpoint of the physical scientist, not the techniques and viewpoint of the lawyer or the major general or even of the engineer.

Operations research has been defined[2] as the application of the techniques of the physical scientist to the study of the operations of war and peace. This may also be only a partial definition; let us see what it would imply. It would mean, for example, that the operations research worker should look at the operation under study with the same impersonal objectivity with which he looks at the problem of the diffraction of light, for example. He must put aside, for the time, questions of usefulness and urgency and must seek to understand what goes on in the sense in which the scientist uses the word understand.

ESTABLISHING A POINT OF VIEW

This very requirement of viewpoint has been one of the greatest sources of friction and misunderstanding in the development of operations research. The executive, whether he is military or industrial, is anxious to have a particular problem solved, so as to help him reach a decision. He finds it difficult to understand why the operations research worker insists on wandering all over the shop, why he does not stick to the point. But in scientific research in general, one does not reach new discoveries by sticking to the point. If Faraday had stuck to the point of trying to develop easily transportable power, he wouldn't have mooned around with such uninteresting things as electric induction. As you probably know, England's prime minister, Gladstone, once asked Faraday what earthly good this electricity would ever be, and Faraday answered that it might eventually be something Mr. Gladstone might be able to tax. Such questions and answers, or their equivalent, are common at the start in any branch of research.

Administrators in general, even the high brass, have resigned themselves to letting the physical scientist putter around with odd ideas and carry out impractical experiments, as long as the things experimented with are solutions or alloys or neutrons or cosmic rays. But when one or more of them start prying into the workings of his own smoothly running organization, asking him and others embarrassing questions not related to the problem he wants them to solve, then there's hell to pay. It takes a great deal of self-restraint on the part of the executive, and requires a great deal of tact on the part of the operations research worker to get safely through this first several months of apparently aimless prying, until the picture begins to come clear to the research worker and his first theories are ready for testing. It says a great deal for the flexibility of our senior military men and of some of our industrialists, that they have been willing to put up with this long enough to get results.

In operations research it isn't possible to shut the worker off in his laboratory and forget about him till he is done. The thing he is studying is, in part, the organization itself and the executive, in a way, is part of the experiment. In fact, the research can only be successful if the operations research group is in frequent, face-to-face contact with the executive in charge of the operation. All the facts about the operations must be available, and usually these cannot be obtained from subordinates. Security restrictions, military or commercial, must not hamper the collection of data. (Believe me when I say that commercial security is just as hampering and just as hard to break through, as is the military brand.)

CURIOSITY UNLIMITED

In some brands of fiction the scientific point-of-view is pictured as cold, impersonal, and somehow basically cruel. I don't believe this is so, or at least that it has to be so. What is needed is a driving curiosity and an almost complete refusal to take anyone's word for anything. The curiosity need not appear cold, and with a bit of tact, the "I'm from Missouri" attitude need not be cruel. Nevertheless, both can be irritating when an organization is being investigated. I can remember the trying time Admiral Low had when we were trying to find out about German submarine tactics in the Caribbean. We suggested trying certain variations of the use of search radar in our antisubmarine planes to see what the U-boat reaction might be. The Admiral finally consented and sent out the order by dispatch, requesting special reports on the results. When the first reports came in, they didn't make sense, and we suggested maybe the variations weren't being done properly. This, of course, was heresy—the Admiral had sent an order, of course it was being carried out! We persisted, however, and were allowed to send a man down to find out himself what was happening. He found that the dispatch had not been understood correctly. By personal interviews he explained the change desired, so we got our data and were able to say that the U-boats' radar listening equipment was less than fifty percent effective. After that experience we tried out dispatches before they were sent; we would ask several naval officers not closely connected with the work to tell what he would do if he got such an order. This procedure was never understood by the Navy, but by its means we avoided a good many misunderstandings.

APPLICATION OF SCIENTIFIC ANALYSIS

This example was meant to illustrate, in a very simple case, the "I'm from Missouri" attitude. But it also brings me to my next point. If operations research is

the application of the methods of physical science to the study of operations, it must involve the peculiar duality of experiment and theory which runs through all such research. The phenomena under study are observed first, observed in a random sort of way, and a simplified and idealized model of the phenomena, amenable to mathematical analysis, is constructed from this first observation, at first a very crude model with some of its properties conveniently vague. If this model is reasonably correct in its main outlines, then certain deductions which may be drawn from it should be verifiable. If these deductions are not borne out by the suggested experiment, the model is scrapped; if they are borne out, the model can be made more detailed and further predictions can be made for further experimental verification, and so on till the theory is well enough checked to turn over to the engineers for use as a law of nature.

The model of a gas as a conglomerate of helter-skelter atoms was a very crude model seventy-five years ago. By now, by the alternating stages of experimental verification and theory elaboration, we can predict the behavior of various gases under a wide variety of conditions.

One does the same thing in operations research. In studying the effectiveness of an actual railway signaling system, one first sets up an ideal system with which to compare the actual one. Never mind what the ideal system would cost, it is supposed to represent the absolutely best possible. If it turns out that the ideal system is much better than the actual one, as far as safety or speed is concerned, then it will pay to study the differences between ideal and actual and to see whether these differences suggest ways of improving the present system. If, however, the actual system is nearly as good as the ideal, then it probably will be best to drop the whole problem and turn to another which will pay off better.

In studying the effect of a change of manufacturing process of the proper location of new factories, one can set up a sort of standard market area, with a specified range of customers, so many taking so much, so many close at hand and others at distances requiring rail shipments, and so on. The relative numbers to put into this model could be taken from previously collected data or it might have to be taken—at first the figures would not have to be very accurate. Next we would wish to manipulate the model to see what would happen if we should require each factory to serve a larger region. How would the shipping services have to be modified, how would the ratio of manufacturing to warehousing facilities have to be changed, when would it pay to have branch warehouses, and so on? Some of the conclusions, based on this very simple model, can perhaps be checked from data collected in the past, some indirect conclusions can possibly be tried out on the present system of factories. In this manner the model may be brought closer to the existing state of things and may eventually be detailed enough to enable predictions to be made from it. In a very similar way a new distribution system for mail-order catalogues was worked out for a large mail-order company. The idealized model, in this case a model of a consumer market, with "mean spans of interest" and various rates of change of frequency of buying, was checked by a number of means until it was complete enough to allow predictions to be made.

NEED FOR STATISTICAL ANALYSIS

Statistical analysis will often be basic in this process of theory and experimental verification. But the statistics are used as tools, not as an end in themselves. Operations research is not a branch of history, it uses past performance in order to predict what might happen in the future under changed circumstances. If operations research is an application of the techniques of the physical sciences to problems of management, it must also be experimental as well as observational. If past operations do not provide a sufficient check on the correctness of a model, some of the minor conclusions of the theory can often be tried out before major changes are made. If the subject of study, for example, is the maintenance and overhaul procedures in a chain of factories, perhaps a new schedule can be tried out in one factory, for a year or so, or perhaps just one type of machine can be experimented with. Often quite small experiments may provide the additional data needed. Often apparently extraneous experiments will extend one's understanding.

The study of automobile traffic flow may provide an example of some of the methods in initial attack on a problem. The retardation of flow at the intersections of major arteries is an important aspect of such a problem. Several methods of study are being carried out. One is the straight statistical one: Data from a number of intersections are being taken of the number of cars passing through the intersection, per traffic-light cycle, to find how this depends on the fraction of right- and left-hand turns, for example. From the resultant results, their similarities and differences, one can begin to form a model of a "standard-car and driver," which may suggest trying a different time or sequence cycle of the traffic light. Results from these tests would suggest further elaboration and so on.

Another mode of attack, to be carried on in parallel, is to study the dynamics of start-stop of a line of traffic. When the light turns green, the whole line of cars does not start as a unit, the first one starts, then the next and so on, a wave of starting travels down the line. Preliminary and very fragmentary observation seems to indicate a remarkable constancy of this wave velocity, it seems to move at about thirty miles an hour along the row of cars.

If further measurements confirm first impressions, it would imply some interrelation between car design and human reaction time which certainly plays an important part in traffic congestion. Such an analysis of

waves of starting is not exactly "sticking to the point" of studying traffic flow through an intersection, but a thorough understanding of its mechanism might well produce results of much greater value than the purely statistical study of flow. It may suggest changes in automobile design as well as in road design.

SEARCH FOR ENEMY SUBMARINES, FOR EXAMPLE

Another example of the interplay of observation, experiment, and theory-making occurs in the problem of the search for enemy submarines by aircraft. It turned out that, of the U-boats sighted in daylight by British aircraft during the summer of 1941, about 40 percent were attacked by the aircraft before they managed to submerge; the other 60 percent managed to get below the surface before the aircraft was in a position to drop its depth charge. This fraction of submarines caught on the surface by the attacking aircraft turned out to be surprisingly constant for a long time during the war. When the United States entered the war, sighting reports on this side of the Atlantic showed that for all of 1942 this fraction was between 40 and 45 percent, even though our type of plane and our training methods differed from those of the British.

This was the first step in the cycle of research. Once this empirical observation had been made, of the constancy of the fraction of U-boats caught on the surface, use could be made of it in planning tactics and in the design of depth-charge fuses. For instance, 40 percent was a large enough fraction of the attacks to make it worth while to equip antisubmarine planes with ordnance which was primarily effective on surfaced submarines. Depth-charge settings were made extra shallow, for instance, and forward firing rockets were installed. Already the research was beginning to pay off. The change to shallow settings on the depth-charges, by itself, increased the number of U-boat sinkings by more than 50 percent.

INTRODUCTION OF PSYCHOLOGICAL QUANTITIES

The next stage in the research was to go behind this empirical fact, to discover why it was that the fraction of submarines caught on the surface remained so constant. In this phase all branches of science were called into play. The laws governing the sighting of a submarine by an aircraft observer depend on the visual contrast between the normal ocean surface and the wake of the submarine, the optical transmission properties of the atmosphere, the methods by which a person's eye scans a large area, the sensitivity of the eye retina to contrast, and the altitude and speed of the plane. Some of these quantities are physical, some physiological; they are all measurable. The sighting of the oncoming aircraft from the submarine depends on a similar set of physical and physiological quantities. The relationship between these two sets of quantities determines the fraction of U-boats which have insufficient time to submerge before the aircraft attacks them.

From this point of view, the statement that 40 percent of the submarines are attacked on the surface corresponds to an interrelation between physical and physiological optics and the physical capabilities of submarines and aircraft. The interesting and important point of the study was that this over-all quantity remained remarkably constant, in spite of wide variations of extraneous circumstances, as long as the equipment involved remained approximately the same. Of course, as soon as the Germans made a radical change in submarine design the fraction mentioned above changed appreciably.

It then was necessary to set up physiological experiments to determine the behavior of the human eye when scanning the horizon. Curiously enough, such experiments had not been done in detail before and the work resulted in important and valuable new data concerning the use of the eye as a search mechanism. This new knowledge could then be employed in the original problem to gain a much broader picture of the mechanism of sighting submarines from aircraft and vice versa. Finally, when the new visual data were combined with optical measurements of the atmosphere and of the properties of the ocean surface, it became possible to put together a fairly complete theory of visual aircraft search, by means of which one could predict what would happen in case the equipment were changed.

At this point, it was possible to design the operation of search by eye so as to make it most effective. It was possible, for instance, to specify the way an observer on the ground should look for an enemy aircraft, which turned out to be different from the way an observer in an antisubmarine plane should look for the submarine wake. Moreover, the results of the measurements could be expressed in mathematical form by use of the calculus of probability. Objects are not usually seen the first time they are looked at. There is a certain probability of actually seeing an object each time the eye glances that way. The calculus of probabilities provides just the kind of analytical machinery to express this behavior.

PROBABILITY FUNCTIONS

In addition, it was seen that the same mathematical techniques, which expressed the probability of visual sighting, could also be used to express the probability of detection by means of radar or sonar. In each case, there was a certain probability function, which depended on the physical or physiological properties of the detection instrument, and which could be measured in the laboratory. This probability function was then a measure of the effectiveness of the search instrument against the object which is searched for; an enemy plane in the sky, an enemy U-boat on the ocean surface, or a submerged U-boat. Actual operational data on number

of sightings could then be checked with the laboratory results to determine operational factors measuring the degree of training of the operator and the average inattentiveness of human beings. These then could all be combined mathematically to give an over-all probability of discovering the target vessel as a function of range, speed, and other properties of the searching vessel.

Finally then, the mathematics of probability could be used to compute the probability of sighting the target for various choices of courses for the searching vessel. When we reached this point in our research, we could say that the theory of search was established and that we were in a position to design optimum search tactics for any situation and optimum equipment for any tactic. For example, a search plan was designed, using the theory by which only two squadrons of planes were able to patrol a band across the whole South Atlantic from Brazil to Africa. This band was "watertight" enough to catch five out of six German blockade runners trying to carry rubber from Japan to occupied France.

This then is the general procedure of operations research. First is the recognition of the constancy of certain quantities typical of the operation, by means of which one can predict the outcome of future operations as long as the equipment is unchanged. Next comes an analysis of the properties of various elements of the operation and perhaps subsidiary laboratory measurements. Out of this comes a fairly detailed theory of the whole operation, a working model, from which one can predict the outcome of future operations even if equipment or conditions be changed. When this stage has been reached, one has obtained a truly scientific insight into the details of the operation and it is possible to design the operation to give the results desired and to advise as to the best equipment and personnel to use. One can envisage exactly the same development in the study of automobile traffic, for example.

WHAT ABOUT QUANTITATIVE ANALYSIS?

A good many attempts at defining operations research include the word "quantitative." If the methods of operations research are those of the physical sciences, they must involve quantitative analysis and they must rely heavily on the tools of mathematics. To be quantitative does not require being exact of course. It often is more fruitful to state that a certain result occurs more than three times and less than ten times oftener than another than it is to state that the ratio between the two frequencies is 4.5738, or whatever. There is an important qualitative difference between the statement, "this quantity is two orders of magnitude larger than that" and the statement "this is very much larger than that."

The fact that the field of study of operations research is highly variable and includes many factors beyond our control certainly does not hamper our use of mathematics. The physical sciences have long ago outgrown rigid determinism. Acoustics might be called an exact science, but it deals with the average effects of clouds of atoms, moving at random. The speed of sound, acoustics' fundamental constant, is only the average resultant of a wide variety of possible atomic interactions. All atomic and nuclear phenomena are random events, often occurring quite infrequently, yet the mathematics of the quantum theory makes their prediction as possible and the theory as tractable as the classical dynamics of planetary orbits. An atom of radium may not disintegrate for thousands of years, but from the mean lifetimes of many such atoms we can reconstruct a model of the nucleus which enables us to predict other reactions. The initiation of an atom-bomb explosion is a highly random process, there is a definite probability that it will be a "dud"; but such considerations need not keep us from computing this probability and designing things to minimize its value. The physical sciences have developed their own brands of mathematics to take into account high degrees of variability; the quantum theory is based on a principle of uncertainty.

TWO ILLUSTRATIVE PROBLEMS

Operations research uses these concepts and these same mathematical techniques to study a new subject matter. The concepts, as usual, are more important and more difficult to master than the techniques. Two simple problems, taken from a course in operations research, will perhaps illustrate what I mean.

Subway express trains go past the center platform at the Penn Station stop of the New York 7th Avenue subway line with equal numbers in each direction each hour of the day. A man arrives on the platform sometime between 9 and 9:30 each weekday morning, having business both uptown and downtown, so he takes whichever directed train comes in first. After a month's routine he finds he takes a downtown train three times as often as he takes an uptown train. What does this result imply concerning the nature and the rigidity of the scheduling of express trains on the 7th Avenue line?

The second problem deals with the installation of aircraft-warning radar on our submarines during the past war. A radar set which could see a Jap plane farther off than a lookout could, would enable our subs to dive in time to escape detection or attack. On the other hand, if the Jap planes had receivers to hear our radar signals, the new radar would increase our danger of detection, not decrease it. Suppose the average range of detection of Jap planes by the new radar is twice the average distance at which a lookout sees a plane. How much more often must planes be sighted when the radar is on than when it is off (and lookouts are used) before one should conclude that Jap planes are hearing our radar sets? We got the right answer to this question during the war, and kept our radar on, thereby pro-

viding adequate warning against Jap planes in most cases. The Germans came to the wrong conclusion, their subs were often caught by surprise by our aircraft and several hundred were thus sunk.

A similar balance of probabilities also turns up in a study of hotel reservations. At any given time a hotel has so many guests in residence and has so many reservations for a week from that day, for instance. How many more reservations should the booking clerk accept? If he turns away all reservations when he has his hotel full on paper, he will never have a full hotel, for about one out of five reservations never show up. On the other hand, he cannot always accept all reservations, for occasionally he will be called on to house more people than the hotel has room for. He must somehow balance the gain in having a larger number of rooms rented against the loss in good will of occasionally failing some one who has been given a reservation.

As soon as the hotel manager has decided how much such a failure costs the hotel, the other factors in the analysis are straightfoward probability analysis. The computation of the mean stay of a guest and its relation to the check-out rate of guests already present, the curve for reservation cancellations, the dependence of fluctuations on time of week and so on; all these can be reduced to a fairly simple set of rules for the booking clerk to follow in making his day by day decisions. One hotel chain is now trying out such a set of rules for a year at one of its hotels to see what bugs develop and to see what improvement results.

OPERATIONS RESEARCH AND COMMUNICATIONS THEORY

Some of the newer applications of mathematics have even more direct connection with the problems of operations research. Communications theory,[3] as developed by Shannon and others, should become a very powerful tool for analyzing the flow of orders in an organization, for example. Even in its simpler aspects the concept of the entropy of an information source and of its communication channel at once makes available the conception machinery of statistical mechanics to such new fields as automatic switching, cryptography, and multi-channel telephony; at the same time it gives these subjects a quantitative base. But they are trivial compared to the possibilities for future application—in the field of language, for example. How much redundancy must there be in a language in order that the meaning to be imparted will survive lapses in attention of the listener? A radar operator sees the screen he is supposed to be watching less than half the time, according to wartime data; we must all of us really receive but a fraction of the words we hear or read. Things must be said twice or thrice over to have a good chance of transmission from mind to mind. Natural languages have this necessary redundancy built into them. Certainly the wording of orders to be understood should be based on a quantitative understanding of redundancy.

The theory of communication enables us to compute the following interesting facts: that the statistical redundancy of the English language is about 50 percent and that for a redundancy larger than this, two-dimensional crossword puzzles are prohibitively difficult to devise, whereas for a redundancy of about 30 percent or less, three-dimensional crossword puzzles should be possible. No doubt such statements appear trivial at first, but a little thought should convince us that they are really very deep and fruitful statements concerning the structure of our written language, and should eventually form the basis for decisions as to how concise one can make orders and still have them understood, for example.

The experiments of Bavelas[4] on the transmission of ideas and orders in task-oriented groups points the way to other applications of communications theory. Bavelas restricts the channels of communication between members of the group to some specified pattern and then observes group action as the members try to solve some cooperative task. Some of the action, the nature and number of the person-to-person messages sent, the emergence of leadership in the group and the frequency of failures, for example, show a definite correlation with the nature of the prescribed communication net. As yet the experiments have not developed to a stage where any but the most rudimentary analysis is required, but further extension of the work to larger groups or more complex tasks will probably need a combination of communication theory and topology to unravel. The combination should throw considerable light on problems of organization, which are now solved (sometimes effectively, sometimes very poorly) by rule-of-thumb.

THEORY OF GAMES

In addition to these new techniques of mathematics, which are shared with other applications of science, operations research has its own branch of mathematics, called the Theory of Games, first developed by von Neumann[5] and since worked on intensively by Project Rand and other military operations research groups. Here the aspects of competition are specifically included in the mathematics. For specific situations, for example, one can now calculate how often one should bluff and how often one should call what appears to be the opponent's bluff.

The cases so far worked out in detail are very elementary ones and the applications of value have been so far chiefly in the field of military tactics. The duel problem is a simple example of the sort of problem amenable to game theory analysis. Each duellist starts for his opponent with one bullet in his pistol. When should he shoot? If he shoots too soon, he will run the risk of missing his opponent, and if he waits too long his opponent may shoot him first.

When the two pistols are identical the problem is simple, but what is the answer when the pistols differ

in range or accuracy? What is the answer when one opponent is the pilot of a jet fighter plane and the other is the tail gunner of a bomber? Airplane battle tactics are now worked out by the use of game theory. One can see that many problems in industrial operations would also be proper subjects for game theory analysis. Any operation involves some elements of competition, either within the organization or in its contacts with other organizations. As game theory is developed, it should be of use in analyzing such elements of the operation.

I have now reviewed a few of the attempts at defining operations research; I will not attempt to give you a definition of my own. I am quite sure that any such attempt would not change the probability of my being asked, tonight or tomorrow, to give a one-sentence definition of the subject and, as I hope I have indicated, I can't think of a good short definition. It is quite possible that it is too early to give a good definition. Chemistry and physics must have been difficult to define in the early days.

OPERATIONS RESEARCH, A PROFESSION?

It is possible that operations research will develop into a well-recognized branch of applied science, with degrees being given in it and with operations research groups being regular parts of any large industrial staff, as they are now of military staffs. It is also possible that the techniques and point of view I have been discussing in this talk will be merged with those of economics and psychology and administration theory to form a larger field, called something else. All I am sure of is that the techniques I have described are making contributions of value and will continue to so do.

It is possible that operations research is but one manifestation of the increased utilization of the physical sciences and of physical scientists in new fields of investigation. A certain amount of research in psychology and psychiatry, carried on the last five years, vaguely indicates that there are considerable differences in the way different scientists think. A detailed psychological analysis of 20 outstanding physicists, 20 biologists, and a corresponding number of psychologists, anthropologists, mathematicians and so on, indicates that the methods of reasoning and the mental reactions to problems differ quite widely between scientists in different fields,[6] differ much less between workers in the same field. Most physicists think in terms of visual images, whereas most psychologists think in terms of spoken or acoustical images, for example. Whether the training in a given science has imposed its way of reasoning on the expert, or whether those who already think in a given way have a better chance of success in a given science is not yet clear; indeed the whole study has so far produced only a few tentative conclusions. Nevertheless, these first results are quite suggestive.

It may just be that our knowledge of psychology and of economics and of management practice has developed to a stage where workers who think the way physical scientists think can be of use in the solution of problems of management. It may be that what is needed is to drain off a certain fraction of these young men who can think this way, who at present are attracted to the physical sciences, and set them to work on the problems involved in running an organization, no matter what the activity is called.

Don't get me wrong, I am *not* saying that the techniques of physical science are any better than any other, nor am I implying that men in this field are any brighter than others. I am not saying or implying that physicists should take over the jobs of the big administrators. I am only raising the point that the techniques and ways of thinking of the physical scientists seem to be of use in *helping* the administrator to do his job better.

SCIENCE ATTEMPTS TO INCREASE UNDERSTANDING

Science and the scientific method have often been criticized for producing machines and weapons and then not seeing that these products are wisely or effectively used. In a rather fundamental sense, however, this criticism involves a contradiction in terms. Science and scientific techniques are not designed to make decisions, to carry on the basic operations of *administration* such as running an army or a business. They have been designed to increase our understanding about some facet of nature, and while this increased understanding will make the task of some administrator wider in scope or will enable him to make his decisions more wisely or will reveal to him more fully the consequences of his decisions, it seldom will *make the decision for him.* The administrative decisions concerning the use of the products of science should be based on all the data available, but they nearly always involve other factors of politics or morale or personalities, which are nonscientific and which must be decided in an arbitrary sort of way. In many such cases it is more important that a decision be made than how it is made, which, of course, is a thoroughly unscientific attitude. It is not surprising, therefore, that the training and temperament of the usual administrator, military or nonmilitary, are quite different from those of the usual scientist. Science requires an attitude of skepticism and an impatience with arbitrary decisions; administration must eventually make arbitrary decisions and often is impatient with criticism. Scientists, as *scientists*, cannot contribute *directly* to the making of decisions concerning the use of the gadgets they have helped create. Some scientists, of course, have turned into very good administrators and it may be that all administrators would be helped by scientific training, but this is beside the point of the present discussion, which is that the process of making an executive decision is not a scientific activity.

But if science is reduced to helping in these decisions only indirectly, is the physical scientist's only contri-

bution a continual production of new gadgets, to further complicate the problem? I believe not. I believe it is possible for a physicist to make an important contribution, as a scientist, in the running of the operations of civilization. He can provide this assistance by studying these operations rather than trying directly to make decisions about them, by helping to understand them as quantitatively and impersonally as possible. Here the object of study is the operation itself; the flow of goods and people in a city, the operation of strategic bombing, the requirements and limitations of a whole communication system are examples. This is not to say that the physicist should not contribute directly, *as a citizen*, to the reaching of decisions concerning these operations, nor is it meant to imply that there is no social value in physics itself; it is to say that the physical scientists has in addition the opportunity (and perhaps the duty) to contribute *as a scientist* to the clarification of the factual basis for these decisions, so that they may be made as wisely as possible.

BIBLIOGRAPHY

1. "Operations Research with Special Reference to Non-military Applications," National Research Council, 1951.
2. P. M. Morse and G. E. Kimball, *Methods of Operations Research* (Technology Press and John Wiley & Sons, Inc., New York, 1950).
3. C. E. Shannon and W. Weaver, *Mathematical Theory of Communication* (University of Illinois Press, 1949).
4. A. Bavelas, J. Acoust. Soc. Am. **22**, 725 (1950); R. F. Bales, *Interaction Process Analysis* (Addison-Wesley Press, Inc., Cambridge).
5. J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton University Press, Princeton, 1944); J. McDonald, *Strategy in Poker, Business and War* (W. W. Norton Company, New York, 1950); *Contributions to the Theory of Games*, Ann. of Math. Studies (Princeton University Press, Princeton, 1950), No. 24.
6. Anne Roe, "A Psychological Study of Physical Scientists," Genetic Psychology Monographs **43**, 121 (1951).

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 2 FEBRUARY, 1952

Contributed Original Research

On the Accumulation of Errors in the Numerical Solution of Differential Equations*

A. PAPOULIS
Moore School, University of Pennsylvania, Philadelphia, Pennsylvania
(Received February 28, 1951)

The following gives a method for the evaluation of errors in the numerical solution of systems of differential equations.

Under certain assumptions the errors can be obtained by finding the difference $r(x)$ between the computed solution $y(x)$ passing through the given initial point, and the solution $Y(x)$ passing through a neighboring point—that is, by solving the same equations repeatedly with different initial conditions.

I. INTRODUCTION

FOR the numerical solution of the differential equation $dy/dx=f(x, y)$ with the initial condition $y(a)=y_0$, the given interval (a, b) is divided into n parts with the points $a=x_0<x_1<\cdots<x_n=b$; and from the value y_{k-1} of the function at $x=x_{k-1}$ and possibly values of y in preceding points, the value y_k of the function at the following point x_k is obtained. In this process the following errors are made:

1. *Truncation error* resulting from the fact that the given function is approximated in the interval (x_{k-1}, x_k) by another, usually a polynomial, passing through the point (x_{k-1}, y_{k-1}); with this approximation the value y_k is obtained from y_{k-1} with a small number of numerical operations. The magnitude of this error can easily be estimated; it depends on the choice of the approximating function.

2. *Round-off-error* resulting from the fact that in computing we use a finite number of decimal places. It can also easily be estimated and it is usually of the order of the last available decimal place.

Let us denote by β_k the difference between the value of the correct solution of our equation at $x=x_k$ which passes through the point (x_{k-1}, y_{k-1}), and the computed value y_k at $x=x_k$; β_k is the sum of the errors 1 and 2.

If we assume that $y(x)$ is correct up to the point x_{k-1} and from x_k to b (correct in the sense that it satisfies the given differential equation), then we shall have committed at $x=b$ an error $E_k=y(b)-y_n$, between the computed y_n and the desired $y(b)$, because of the error at $x=x_k$.

Under the assumption A, as will presently be stated, the total error R resulting from the errors in all steps is the sum of the errors E_k of each step. Thus knowledge of the E_k's enable us to find the total error R at $x=b$ resulting from all the errors at each point x_k.

For the determination of the E_k's we first find $r(x)=Y(x)-y(x)$ where $Y(x)$ is a solution of (1) with the initial condition $Y(x_0)=y_0+M$, where M is a constant; then it is shown that $E_k=\beta_k[r(b)/r(x_k)]$ and hence

$$R=r(b)\sum_{k=1}^{n}\frac{\beta_k}{r(x_k)}.$$

The quantity $r(x)$ satisfies Eq. (3), and it can be obtained by approximately solving this equation.

But under assumption B, as stated later, $r(x)$ can be obtained by determining $y(x)$ and $Y(x)$, that is, by solving (1) twice with different initial conditions. Thus solution of auxiliary equations for the determination of the total error R is avoided.

The method also applies for a system of more than one equation. We shall first discuss the simple case (1) to illustrate the method.

II. THE EQUATION OF THE FIRST ORDER

$$dy/dx=f(x, y). \tag{1}$$

Suppose $y(x)$ and $Y(x)$ are two solutions of (1); with $r(x)=Y(x)-y(x)$, we have

$$[dr(x)/dx]=f(x, Y)-f(x, y). \tag{2}$$

Assumption A. With $r(x)=Y(x)-y(x)$, $f(x, Y(x))-f(x, y(x))=r(x)\varphi(x)$, where $\varphi(x)=(\partial/\partial x)f(x, y(x)+\theta(x)r(x))0\leq\theta<1$, we assume $r(x)$ small compared to $y(x)$; then $\varphi(x)\simeq(\partial/\partial x)f(x, y(x))$.

Thus we have from (2)

$$\frac{dr(x)}{dx}=\varphi(x)r(x) \tag{3}$$

$$r(x)=C\exp\left(\int_{x_0}^{x}\varphi(x)dx\right) \tag{4}$$

where C is a constant.

Suppose that $r_1(x)$ and $r_2(x)$ are two different solutions of (3); then as it can be seen from (4)

$$r_2(x)=kr_1(x), \tag{5}$$

* This work and preparation of this paper was sponsored by the Burroughs Adding Machine Company.

where

$$k=[r_2(x_0)/r_1(x_0)].$$

In the following by $y(x)$ we shall mean the desired solution of (1) with the initial condition $y(x_0)=y_0$, and by $r(x)$ the difference $Y(x)-y(x)$ where $Y(x)$ is another solution of (1) such that assumption A is satisfied. Thus $r(x)$ is a solution of (3).

Suppose that $Y_1(x)$ is a solution of (1) such that $Y_1(x_1)=y(x_1)+\beta_1$; then, since $Y(x)-y(x)$ is a solution of (3), it follows from (5) that

$$Y_1(x)-y(x)=kr(x). \tag{6}$$

If we replace x by x_1 in (6), we obtain

$$k=[\beta_1/r(x_1)];$$

thus

$$Y(x)-y(x)=[\beta_1 r(x)/r(x_1)]. \tag{7}$$

Similarly, with $Y_k(x)$ a solution of (1) such that

$$Y_k(x_k)=Y_{k-1}(x)+\beta_k, \tag{8}$$

we must have

$$Y_k(x)-Y_{k-1}(x)=[\beta_k r(x)/r(x_k)]. \tag{9}$$

This being true for every k, we have after h such relations

$$Y_h(x)=y(x)+r(x)\sum_{k=1}^{h}\frac{\beta_k}{r(x_k)}. \tag{10}$$

Suppose that $Y_{k-1}(x)$ is the true solution of (1) in the interval (x_{k-1}, x_k) whose value at the beginning of this interval equals the computed value y_{k-1} of y at this point. In computing $Y_{k-1}(x_k)$ an error β_k is committed so that the true solution $Y_k(x)$ of (1) in the following interval satisfies (8) in the beginning of that interval. Thus from the preceding material it follows that the computed value of y at $x=x_h$ is related to the true value $y(x_h)$

$$Y_h(x_h)=y(x_h)+r(x_h)\sum_{k=1}^{h}\frac{\beta_k}{r(x_k)},$$

and for $h=n$ we have with $R=y_n-y(x_n)$ the total error at the point $x_n=b$,

$$R=r(b)\sum_{k=1}^{n}\frac{\beta_k}{r(x_k)}. \tag{11}$$

Relation (11) gives us the total error R by means of $r(x)$ and of the errors β_k at the different steps.

The sum can be used to give a statistical or maximum value of the error. If n is large, it can easily be changed into an integral whose approximate value can be obtained with a number of steps much smaller than n. Indeed, with $(b-a)/n=\Delta x$, (11) becomes

$$R=\frac{r(b)}{\Delta x}\int_a^b\frac{\beta(x)}{r(x)}dx, \tag{11a}$$

where $\beta(x)$ is the value of the error committed at the step near x. If the errors are only the result of round-off and if they are random with the same dispersion d^2, then (11a) gives for the dispersion of R

$$\sigma^2(R)=\frac{d^2r^2(b)}{\Delta x}\int_a^b\frac{dx}{r^2(x)}. \tag{11b}$$

$r(x)$ can be obtained either from (4), that is by solving (3), or by the following method, if we can satisfy assumption B.

Assumption B. We can find our $r(x)$ satisfying assumption A, but such that $r(x_k)\gg R_k$, where $R_k=y_k-y(x_k)$ is the total error at the point x_k.

With this assumption satisfied, we find numerically a solution of (1) with initial condition $Y(x_0)=y_0+M$, where M to be so chosen that $Y(x)-y(x)$ satisfies assumption B, then

$$r(x)=Y(x)-y(x)$$

is the desired function.

Assumption B is necessary because $Y(x)$ and $y(x)$ are computed numerically, and in order to have a satisfactory estimate of $r(x)$, this difference must be large compared to the error committed in their evaluation.

If the magnitude of $r(x)$ increases with increasing x so that assumption A is not satisfied or decreases so that assumption B is not satisfied, we can divide the given interval (a, b) into a suitable number of pieces, obtain in each of them one $r_m(x)$, and then combine them into one $r(x)$, changing their scale since they are proportional (see (5)).

III. THE SYSTEM OF TWO EQUATIONS

$$\begin{aligned}(dx/dt)&=f(x, y, t),\\ (dy/dt)&=g(x, y, t).\end{aligned} \tag{12}$$

Suppose that $[x(t), y(t)]$ and $[X(t), Y(t)]$ are two solutions of (13); with $r(t)=X(t)-x(t)$, $s(t)=Y(t)-y(t)$ we have

$$\begin{aligned}(dr/dt)&=f(X, Y, t)-f(x, y, t),\\ (ds/dt)&=g(X, Y, t)-g(x, y, t).\end{aligned} \tag{13}$$

Assumption A. $[X(t), Y(t)]$ is close enough to $[x(t), y(t)]$ so as to neglect the higher powers of $r(t)$ and $s(t)$ in the series-expansion of $f(X, Y, t)-f(x, y, t)$ and $g(X, Y, t)-g(x, y, t)$.

Under this assumption we obtain from (13)

$$\begin{aligned}(dr/dt)&=(\partial/\partial x)f(x, y, t)r(t)+(\partial/\partial y)f(x, y, t)s(t),\\ (ds/dt)&=(\partial/\partial x)g(x, y, t)r(t)+(\partial/\partial y)g(x, y, t)s(t).\end{aligned} \tag{14}$$

Suppose that $[r_1(t), s_1(t)]$ and $[r_2(t), s_2(t)]$ are two solutions of (14); then, with k and h two arbitrary constants, $[kr_1(t)+hr_2(t)]$, $[ks_1(t)+hs_2(t)]$ must also be

a solution. Indeed,

$$\frac{d(kr_1+hr_2)}{dt}=k\frac{dr_1}{dt}+h\frac{dr_2}{dt}=k\left(\frac{\partial f}{\partial x}r_1+\frac{\partial f}{\partial y}s_1\right)$$

$$+h\left(\frac{\partial f}{\partial x}r_2+\frac{\partial f}{\partial y}s_2\right)=\frac{\partial f}{\partial x}(kr_1+hr_2)+\frac{\partial f}{\partial y}(ks_1+hs_2).$$

Similarly for ks_1+hs_2; we shall say that the two solutions $[r_1(t), s_1(t)]$ and $[r_2(t), s_2(t)]$ are linearly independent if with

$$D(t)\equiv r_1(t)s_2(t)-r_2(t)s_1(t), \tag{15}$$

we have

$$D(t)\not\equiv 0 \text{ for every } t. \tag{16}$$

If (16) is not true for one value of t, say $t=t_k$, that is if $D(t_k)=0$, then $D(t)=0$ for every t.

Conversely, suppose that $[r_1(t), s_1(t)]$ and $[r_2(t), s_2(t)]$ are two independent solutions of (14), and $[r(t), s(t)]$ a third solution. Then two constants k and h can be determined such that

$$\begin{aligned} kr_1(t)+hr_2(t)&\equiv r(t),\\ ks_1(t)+hs_2(t)&\equiv s(t). \end{aligned} \tag{17}$$

The foregoing statements follow from well-known properties of fundamental systems of solutions of linear equations.

In the following by $[x(t), y(t)]$ we shall mean the desired solution of (12) with the initial conditions $x(t_0)=x_0$, $y(t_0)=y_0$, and by $[r_1(t), s_1(t)]$ and $[r_2(t), s_2(t)]$ the differences $[X'(t)-x(t), Y'(t)-y(t)]$ and $[X''(t)-x(t), Y''(t)-y(t)]$, respectively, where

$$[X'(t), Y'(t)], \quad [X''(t), Y''(t)]$$

are solutions of (12) such that assumption A is satisfied, and $[r_1(t), s_1(t)]$, $[r_2(t), s_2(t)]$ are linearly independent; thus $[r_1(t), s_1(t)]$ and $[r_2(t), s_2(t)]$ are solutions of (14). $D(t)$ we define as in (15).

Suppose that $[X_1(t), Y_1(t)]$ is a solution of (12) such that

$$X_1(t_1)=x(t_1)+\delta_1 \quad Y_1(t_1)=y(t_1)+\beta_1, \tag{18}$$

then, since $[X_1(t)-x(t), Y_1(t)-y(t)]$ is a solution of (14), we must have (see (17))

$$\begin{aligned} X_1(t)-x(t)&=kr_1(t)+hr_2(t),\\ Y_1(t)-x(t)&=ks_1(t)+hs_2(t). \end{aligned} \tag{19}$$

If we give to t the value t_1 in (19) and solve the system for k and h using (18), we obtain

$$k=\frac{\delta_1 s_2(t_1)-\beta_1 r_2(t_1)}{D(t_1)},$$

$$h=\frac{-\delta_1 s_1(t_1)+\beta_1 r_1(t_1)}{D(t_1)}.$$

Similarly with $[X_k(t), Y_k(t)]$ a solution of (12) such that

$$X_k(t_k)-X_{k-1}(t_k)=\delta_k, \quad Y_k(t_k)-Y_{k-1}(t_k)=\beta_k, \tag{21}$$

we must have

$$\begin{cases} X_k(t)=X_{k-1}(t)+\dfrac{\delta_k s_2(t_k)-\beta_k r_2(t_k)}{D(t_k)}r_1(t)\\ \qquad\qquad +\dfrac{-\delta_k s_1(t_k)+\beta_k r_1(t_k)}{D(t_k)}r_2(t),\\ Y_k(t)=Y_{k-1}(t)+\dfrac{\delta_k s_2(t_k)-\beta_k r_1(t_k)}{D(t_k)}s_1(t)\\ \qquad\qquad +\dfrac{-\delta_k s_1(t_k)+\beta_k r_1(t_k)}{D(t_k)}s_2(t); \end{cases} \tag{22}$$

this being true for every k, we have after m such relations

$$\begin{cases} X_m(t)=x(t)+r_1(t)\sum_{k=1}^{m}\dfrac{\delta_k s_2(t_k)-\beta_k r_2(t_k)}{D(t_k)}\\ \qquad\qquad +r_2(t)\sum_{k=1}^{m}\dfrac{-\delta_k s_1(t_k)+\beta_k r_1(t_k)}{D(t_k)},\\ Y_m(t)=y(t)+s_1(t)\sum_{k=1}^{m}\dfrac{\delta_k s_2(t_k)-\beta_k r_2(t_k)}{D(t_k)}\\ \qquad\qquad -s_2(t)\sum_{k=1}^{m}\dfrac{-\delta_k s_1(t_k)+\beta_k r_1(t_k)}{D(t_k)}. \end{cases} \tag{23}$$

Suppose that $[X_{k-1}(t), Y_{k-1}(t)]$ is the true solution of (12) in the interval (t_{k-1}, t_k), whose value in the beginning of this interval equals the computed value (x_{k-1}, y_{k-1}) at this point. In computing $[X_{k-1}(t_k), Y_{k-1}(t_k)]$ an error (δ_k, β_k) is committed, so that the true solution $[X_k(t), Y_k(t)]$ in the following interval satisfies (21) in the beginning of that interval. From the foregoing it follows that the computed value of $[x(t), y(t)]$ at $t=t_m$ is related to its true value $[x(t_m), y(t_m)]$ by

$$X_m(t_m)=x(t_m)+r_1(t_m)\sum_{k=1}^{m}\frac{\delta_k s_2(t_k)-\beta_k r_2(t_k)}{D(t_k)}$$

$$+r_2(t_m)\sum_{k=1}^{m}\frac{\delta_k s_1(t_k)+\beta_1 r_1(t_k)}{D(t_\kappa)},$$

$$Y_m(t_m)=y(t_m)+s_1(t_m)\sum_{k=1}^{m}\frac{\delta_k s_2(t_k)-\beta_k r_2(t_k)}{D(t_k)}$$

$$+s_2(t_m)\sum_{k=1}^{m}\frac{-\delta_k s_1(t_k)+\beta_1 r_1(t_k)}{D(t_k)},$$

and for $m=n$ we have with $R=X_n-x(t_n)$, $S=Y_n-y(t_n)$ the total errors at the point $t_n=b$.

$$R=r_1(b)\sum_{k=1}^{n}\frac{\delta_k s_2(t_k)-\beta_k r_2(t_k)}{D(t_k)}+r_2(b)\sum_{k=1}^{n}\frac{-\delta_k s_1(t_k)+\beta_k r_1(t_k)}{D(t_k)},$$
$$S=s_1(b)\sum_{k=1}^{n}\frac{\delta_k s_2(t_k)-\beta_k r_2(t_k)}{D(t_k)}+s_2(b)\sum_{k=1}^{n}\frac{-\delta_k s_1(t_k)+\beta_k r_1(t_k)}{D(t_k)}, \tag{24}$$

(24) gives us the total errors (R, S) by means of (r_1, s_1) and (r_2, s_2) and of the errors (δ_k, β_k) at the different steps.

The sum can be used to give a statistical or maximum value of the error. For large n it can be written as an integral easy to evaluate approximately; with $(b-a)/n=\Delta t$, (24) becomes

$$R=\frac{r_1(b)}{\Delta t}\int_a^b\frac{\delta(t)s_2(t)-\beta(t)r_2(t)}{D(t)}dt+\frac{r_2(b)}{\Delta t}\int_a^b\frac{-\delta(t)s_1(t)+\beta(t)r_2(t)}{D(t)}dt,$$
$$S=\frac{s_1(b)}{\Delta t}\int_a^b\frac{\delta(t)s_2(t)-\beta(t)r_2(t)}{D(t)}dt+\frac{s_2(t)}{\Delta t}\int_a^b\frac{-\delta(t)s_1(t)+\beta(t)r_1(t)}{D(t)}dt, \tag{24a}$$

where $[\delta(t), \beta(t)]$ the value of the errors committed at the step near t.

For random error, we obtain expressions similar to (11b).

The functions $[r_1(t), s_1(t)]$ and $[r_2(t), s_2(t)]$ can be obtained either by solving approximately (14) or by the following method, if we can satisfy assumption B.

Assumption B. We can find two solutions $[r_1(t), s_1(t)]$ and $[r_2(t), s_2(t)]$ satisfying assumption A but such that $r_1(t_k), r_2(t_k) \gg R_k s_1(t_k), s_1(t_k), s_2(t_k) \gg S_k$, where $R_k=X_k-x(t_k)$, $S_k=y_k-y(t_k)$ are the total errors at the point t_k.

With this assumption satisfied we find numerically the solutions $[X'(t), Y'(t)]$ and $[X''(t), Y''(t)]$ of (12) with the initial conditions

$$X'(t_0)=x_0+M_1 \quad Y'(t_0)=y_0+N_1$$
$$X''(t_0)=x_0+M_2 \quad Y''(t_0)=y_0+N_2,$$

where (M_1, N_1) and (M_2, N_2) are so chosen that (X', Y') and (X'', Y'') satisfy assumption B and further $M_1/N_1 \neq M_2/N_2$; then the functions

$$r_1(t)=X'(t)-x(t) \quad s_1(t)=Y'(t)-y(t)$$
$$r_2(t)=X''(t)-x(t) \quad s_2(t)=Y''(t)-y(t)$$

satisfy our requirements.

The total errors in (11) and (24) are given in terms of the β_k's; they apply to any step by step method of solution, but do not show how to calculate the quantities β_k, which depend on the specific method used.

The foregoing method is similar to the one developed by Professor H. Rademacher,[1-3] for the Moore School of Electrical Engineering, University of Pennsylvania.

In this paper two auxiliary functions $\lambda(t)$ and $\mu(t)$ are introduced satisfying the "adjoint equations" similar to Eq. (14). The expressions for the errors R and S are similar to (24a). For their determination, knowledge of these functions is necessary, and hence the need of solving a system of two new equations of different form than the given ones.

In the previous discussion the functions $r(t)$ and $s(t)$ can be determined simply by solving repeatedly the given Eq. (12) with different initial conditions, under assumption B.

The foregoing can also simply be extended to the solution of more than two equations; in the case of three equations it will be necessary to determine three triplets of functions $[r_k(t), s_k(t), w_k(t)]$ with $k=1, 2, 3$, and these functions can again be obtained by solving the given system of equations with 3 new initial conditions. The expressions for the final errors are similar to (24) and (24a).

[1] Hans A. Rademacher, "On the Accumulation of Errors in Process of Integration of High-Speed Calculating Machines," Proceedings of a Symposium on Large Scale Digital Calculating Machinery (1948), pp. 176–187. Harvard Computational Laboratory Series, Vol. 16.

[2] F. Schlesinger, Astronom. J. **30**, 183–190 (1917).

[3] Dirk Brouwer, Astronom. J. **46**, 149–153 (1937).

The Thermal Conductivity of Metals at High Temperature*

C. L. HOGAN† AND R. B. SAWYER
Lehigh University, Bethlehem, Pennsylvania
(Received April 4, 1951)

The Forbes bar method has been modified and adapted to the problem of measuring the thermal conductivity of metals and alloys in the temperature range from 25°C to 1000°C. A mathematical analysis is presented which indicates the error involved by assuming plane isothermals in the sample. The electrical and thermal conductivities of commercial "A" nickel, inconel, several stainless steel alloys, and 1010 steel are reported, and the results are compared, qualitatively at least, with the theory presented by Wilson[1] and Makinson.[2]

EXPERIMENTAL PROCEDURE

THE material whose thermal conductivity is to be measured is prepared as a cylindrical sample 0.375 cm in radius and 30 cm long. Chromel-alumel thermocouples of A. W. G. No. 24 wire are tightly peened into small holes at 2 cm intervals along the sample. A small heater (about 2 cm axial length) is wound on the top end of the specimen using A. W. G. No. 26 asbestos-covered Nichrome wire, and No. 12 copper wires are brazed to each end of the sample. Figure 1 indicates the position of these various wires. Before placing the sample in the furnace, the thermocouple wires are placed in fine quartz tubes, and the entire sample is then covered with a thin layer of Alundum cement. The sample is then centered in a cylindrical furnace as shown in Fig. 2. The thermocouple wires are led out of the bottom of the furnace and the Nichrome heater wires out the top. A reference thermocouple is placed in the furnace below the sample and usually in contact with the heavy-walled copper tube which forms the inside lining of the furnace. The furnace tube is then packed with screened Sil-O-Cel.

All thermocouples are connected to a switchboard so that each thermocouple on the sample can be connected in turn differentially with the reference and the resulting emf connected across a Leeds and Northrup type R galvanometer. The deflection of the galvanometer is thus a measure of the difference in temperature between the reference thermocouple and the particular location on the sample where the second thermocouple is attached.

The furnace construction is clearly indicated in Fig. (2).

The furnace is brought to any desired temperature by means of the chromel furnace winding. The furnace voltage is controlled by means of two Sola constant voltage transformers. When a steady state condition is reached, a current of 5–10 amperes dc is passed through the specimen by means of the copper wires brazed to each end. The potential drop per centimeter is measured by connecting the chromel-p wires of each pair of thermocouples in turn to a potentiometer. When the specimen comes to equilibrium with the furnace, the thermocouple in the center of the 30-cm long specimen is connected in differential with the reference, and the

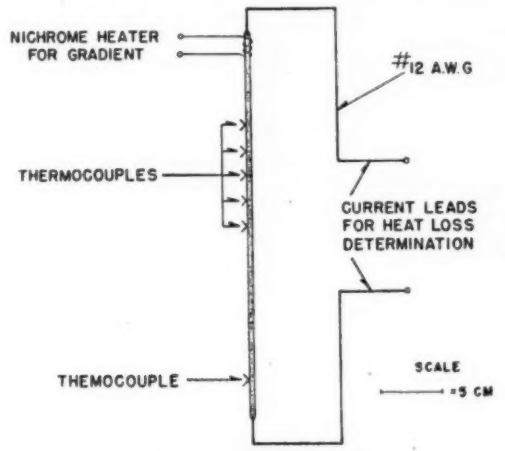


FIG. 1. Specimen showing location of heating coils and thermocouples.

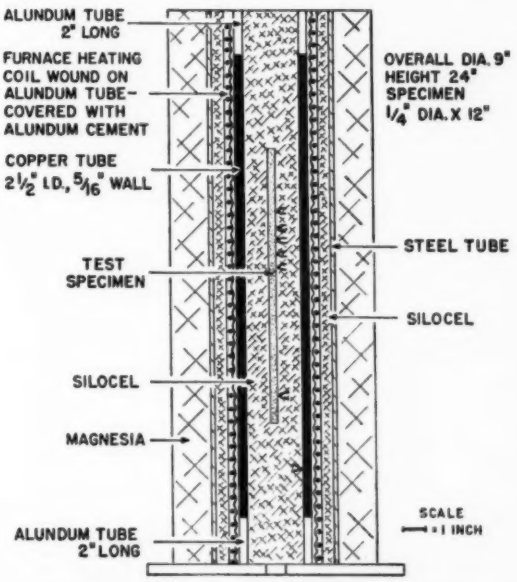


FIG. 2. General construction of furnace showing location of specimen.

* A summary of a dissertation presented by C. L. Hogan to the Graduate Faculty of Lehigh University in candidacy for the degree of Doctor of Philosophy.

† Now at Bell Telephone Laboratories, Murray Hill, New Jersey.

[1] A. H. Wilson, Proc. Cambridge Phil. Soc. **33**, 371 (1937).

[2] R. E. B. Makinson, Proc. Cambridge Phil. Soc. **34**, 474 (1938)

TABLE I. Analyses of specimens.

	C	Mn	Si	Ni	Cr	P	S	Cu	N	Al	Cb	Fe
Inconel X	0.03	0.47	0.39	73.19	14.38	···	0.007	0.03	···	0.83	···	6.99
Nickel	0.06	0.22	0.02	99.48	···	···	0.005	0.05	···	···	···	0.14
Type 303	0.17	0.61	0.51	8.97	18.42	···	···	···	···	···	···	Bal.
Type 310	0.10	1.83	0.84	20.68	25.54	0.025	0.005	···	···	···	···	50.98
Type 347	0.069	1.80	0.70	11.20	18.00	0.021	0.007	···	···	···	0.77	Bal.
Type 446	0.13	0.56	0.50	0.10	26.00	0.012	0.007	···	0.14	···	···	Bal.
1010 steel												

galvanometer is connected in series with them. The galvanometer deflection is then proportional to the temperature difference between the specimen and the furnace lining. The specimen is raised to a temperature of about 1°C above the furnace temperature in this part of the experiment.

The power input per centimeter along the sample is given by

$$\frac{VI}{JL}=\frac{2\pi k_s\Delta T}{\ln(r_2/r_1)},$$

where: V = potential drop (volts) along rod between points L cm apart, I = current in rod (amperes), J = Joule's constant, k_s = thermal conductivity of furnace packing, ΔT = temperature difference between specimen and copper furnace lining when specimen is in equilibrium carrying a current of I amperes (°C), r_2 = radius of copper tube lining furnace, and r_1 = radius of specimen.

It is not necessary that all these constants be measured explicitly. If we let

$$\frac{k_s}{\ln(r_2/r_1)}=r_1H,$$

then H will have the dimensions (calories per second per cm² per degree) and the significance of a heat transfer coefficient, and the equation becomes

$$VI/JL=2\pi r_1H\Delta T.$$

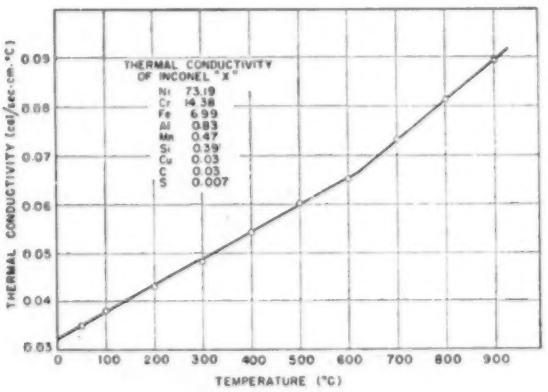


FIG. 3. Thermal conductivity of inconel "X" *versus* temperature.

Further, if we let

$$R=2H/r_1,$$

then

$$R=VI/LJ\pi r_1^2\Delta T.$$

After the preceding data (V/L, I, ΔT) are taken, the current through the cylindrical sample is turned off, and the sample again returns to furnace temperature. At this time a current of approximately 0.5 ampere is passed through the nichrome heater at the top of the sample. When the steady state is attained, the temperature gradient existing along the sample is measured by the method indicated above. The top of the sample is never raised to a temperature of more than 5°C above the bottom during this part of the experiment.

The temperature (recorded as cm of galvanometer deflection) of each thermocouple is then plotted against distance down the sample. On semi-log paper this plot is a straight line. If we denote the slope of this line as m, then the thermal conductivity of the sample is:

$$k=R/m^2.$$

MATHEMATICAL ANALYSIS OF EXPERIMENT

The problem is essentially that of determining the steady state temperature in a semi-infinite rod if one end of the rod is maintained at temperature T_0 and the rod loses heat through its surface at a rate proportional to the temperature difference between the surface of the rod and the furnace. (Because of its large heat capacity, the heavy-walled copper tube used as a furnace liner never measurably changes its temperature during the experiment.)

TABLE II. Thermal conductivity (calories/cm sec °C).

Temperature °C	Inconel X	Nickel	303	310	347	446	1010 steel
0	0.032	0.174	0.034	0.031	0.046	0.054	0.173
100	0.038	0.156	0.036	0.034	0.049	0.057	0.163
200	0.043	0.138	0.038	0.039	0.053	0.061	0.153
300	0.048	0.120	0.042	0.044	0.056	0.065	0.144
400	0.054	0.114	0.047	0.048	0.061	0.068	0.136
500	0.060	0.121	0.051	0.054	0.066	0.072	0.130
600	0.065	0.128	0.055	0.060	0.070	0.076	0.123
700	0.073	0.136	0.059	0.066	0.075	0.080	0.117
800	0.081	0.143	0.064	0.072	0.081	0.084	0.111
900	0.089	0.150	0.068	0.078	0.087	0.087	···

The heat flow problem can be stated as:

$$\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}+\frac{\partial^2 u}{\partial Z^2}=0, \tag{1}$$

$$u(r, o)=T_0, \tag{2}$$

$$\lim_{Z\to\infty} u(r, Z)=0, \tag{3}$$

$$2\pi Ck\frac{\partial u(C, Z)}{\partial r}=-2\pi CHu, \tag{4}$$

or

$$C\frac{\partial u(C, Z)}{\partial r}=-\frac{CH}{k}u(C, Z),$$

or

$$C\partial u(C, Z)/\partial r=-hu(C, Z),$$

where $h=CH/k$, where C=radius of rod, k=thermal conductivity, H=heat transfer coefficient, $u(r, Z)$ =temperature function, and T_0=temperature at $Z=0$.

TABLE III. Electrical conductivity (mho/cm)$\times 10^{-4}$.

Temperature °C	Inconel X	Nickel	303	310	347	446	1010 steel
0	···	11.3	1.38	···	1.27	1.75	···
100	0.814	7.56	1.25	0.999	1.18	1.44	4.30
200	0.804	5.12	1.13	0.951	1.11	1.25	2.82
300	0.794	3.87	1.05	0.908	1.07	1.12	2.05
400	0.784	2.87	0.988	0.870	1.03	1.02	1.60
500	0.774	2.62	0.940	0.841	1.00	0.940	1.20
600	0.764	2.40	0.900	0.818	0.984	0.880	0.95
700	0.770	2.25	0.870	0.800	0.967	0.840	0.80
800	0.782	2.10	0.846	0.785	0.950	0.820	0.70
900	0.795	2.02	0.824	0.771	0.933	0.810	···

This problem can be solved easily. The solution is:

$$u(r, Z)=\sum_{j=1}^{\infty} A_j J_0(\alpha_j r)\exp(-\alpha_j Z),$$

where

$$A_j=\frac{2\alpha_j CT_0 J_1(\alpha_j C)}{(\alpha_j^2 C^2+h^2)[J_0(\alpha_j C)]^2}$$

and α_j are the positive roots of

$$C\alpha_j J_0'(\alpha_j C)=-hJ_0(\alpha_j C). \tag{5}$$

It has been experimentally determined that:

$$h=(3.024\times 10^{-4})/k.$$

Hence the roots of the foregoing Eq. (5) depend upon the thermal conductivity of the sample. The lowest conductivity measured was 0.03 (calories/sec cm °C). For a rod 0.3175-cm radius whose thermal conductivity is 0.03 cal/sec cm °C, the roots of Eq. (5) are:

$\alpha_1=0.4467$ $\quad$ $\alpha_3=22.11$

$\alpha_2=12.07$ $\quad$ $\alpha_4=32.05$, etc.

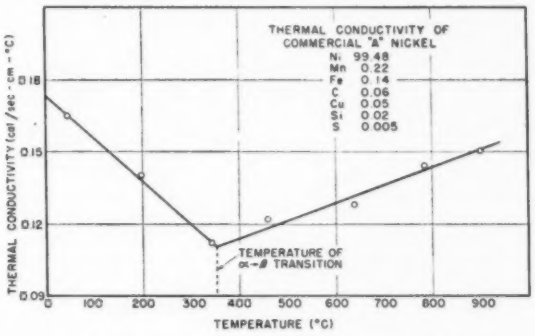


FIG. 4. Thermal conductivity of commercial "A" nickel *versus* temperature.

For the previous conditions, then, the steady state temperature function is:

$$u(0.3175, Z)=0.997T_0\exp(-0.4467Z)+\sum_{j=2}^{\infty}(RT_0/k\alpha_j)\exp(-\alpha_j Z).$$

This series converges so rapidly that only the first term need be considered. Hence

$$u(0.3175, Z)\approx 0.997T_0\exp(-\alpha_1 Z)\approx T_0\exp(-\alpha_1 Z).$$

It is easily seen from this solution that a rod 30-cm long is essentially semi-infinite, provided the thermal conductivity does not exceed 0.25 calories/sec cm °C.

Since the series for the temperature function converges so rapidly that only the first root of Eq. (5) is required, it is possible to find an expression for this root in terms of the thermal conductivity of the sample.

This can be shown to be

$$\alpha_1=(R/k)^{\frac{1}{2}},$$

where

$$R=2H/C=2kh/C^2.$$

Thus the temperature function can be written as

$$u(0.3175, Z)=T_0\exp[-Z(R/k)^{\frac{1}{2}}],$$

which is the exact solution for the problem of a thin

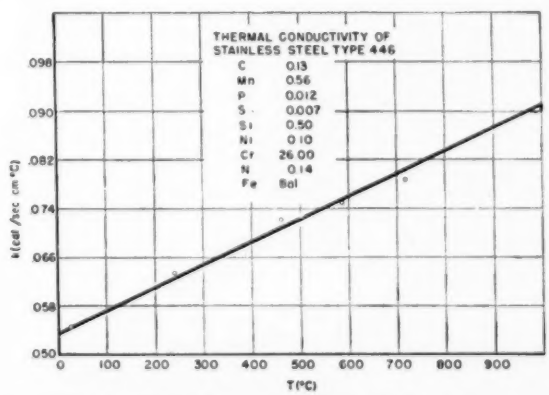


FIG. 5. Thermal conductivity of stainless steel type 446 *versus* temperature.

wire losing heat to its surroundings and with one end maintained at a temperature T_0. It represents the temperature distribution, with an error of less than 1 percent, for a rod 0.3175 cm in radius whose thermal conductivity lies between 0.03 and 0.25 calories/cm °C sec.

RESULTS

There are reported herewith, the values of the thermal conductivity and electrical conductivity of inconel X, commercial "A" nickel, stainless steel alloys type 303, type 310, type 347, and type 446, and 1010 steel. The analyses of these specimens are given in Table I.

The values of the thermal conductivity of these alloys are tabulated in Table II.

The values of the electrical conductivity of these alloys are listed in Table III.

The values of the thermal conductivity of inconel, nickel, and type 446 stainless steel are shown graphically in Figs. 3, 4, and 5.

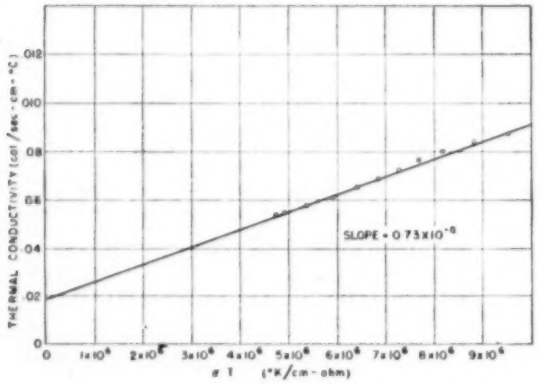


FIG. 6. Lorentz plot (k vs σT) for inconel "X".

Smith and Palmer,[3] after having investigated the electrical and thermal conductivities of a number of copper alloys, found that the thermal conductivity could be represented by a straight line relationship of the form:

$$k = L\sigma T + a,$$

where L and a were constants for a particular set of alloys. This relation has been found to hold for all metals reported on in this paper. As examples, Figs. 6 and 7 show the result of plotting k vs σT for inconel and type 446 stainless.

As pointed out in Makinson's paper[2] the thermal conductivity of metal can be considered to be made up of two terms: $k = k_e + k_g$, where k_e represents the heat transported per unit time by the conduction electrons and k_g represents the heat transported by the lattice vibrations. If we define L_e such that $L_e = k_e/T$; then following Wilson[1] it can be shown that for iron L_e should be always less than the classical Lorentz constant $L_0 = [\pi^2/3][(k/e)^2]$ and L_e should approach L_0 for

[3] C. St. A. Smith and E. W. Palmer, Am. Inst. Mech. Engrs. Trans. **117**, 225 (1935).

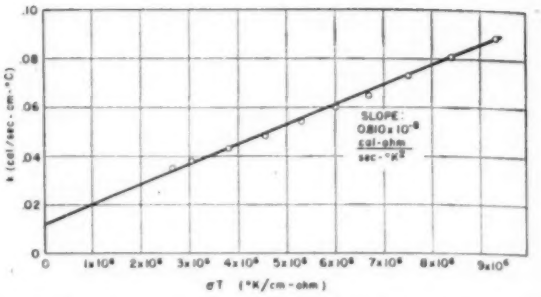


FIG. 7. Lorentz plot (k vs σT) for stainless steel type 446.

$T/\theta > 1$, where θ is the Debye characteristic temperature. The addition of impurities to iron should, when present in small amounts, tend to make L_e approach L_0 more rapidly. It is a question as to how L_e should behave for alloys such as stainless steel and inconel, but it is not unreasonable to assume from Makinson's work that

$$L_e \approx L_0 \quad \text{for} \quad T/\theta > 1,$$
$$L_0 = 0.58 \times 10^{-8}.$$

Thus we should expect for type 446 stainless, for instance, that

$$k = 0.585 \times 10^{-8}\sigma T + k_g.$$

Empirically it is found from Fig. 7 that for this alloy

$$k = 0.73 \times 10^{-8}\sigma T + 0.018,$$

which can be written as

$$k = 0.585 \times 10^{-8}\sigma T + 0.14 \times 10^{-8}\sigma T + 0.018.$$

For the temperature region from 0°C to 1000°C the electrical conductivity of type 446 stainless can be very closely approximated by:

$$\sigma T = 0.51 \times 10^{4} T + 3.36 \times 10^{6}.$$

Hence our empirical equation linking the thermal and electrical conductivities can be written as:

$$k = 0.585 \times 10^{-8}\sigma T + 0.023 + 0.071 \times 10^{-4} T,$$

where we can identify

$$k_e = 0.585 \times 10^{-8}\sigma T = \frac{\pi^2}{3}\left(\frac{k}{e}\right)^2 \sigma T$$

and

$$k_g = 0.023 + 0.071 \times 10^{-4} T.$$

Similar calculations can be carried out for all the alloys reported and reasonable values for the lattice conductivity of all are obtained.

ACKNOWLEDGMENT

The authors wish to express their appreciation to the Bell Aircraft Corporation who furnished all the samples reported on in this paper, and in addition supplied generous grants of funds in order to make possible this work.

Cathodic Sputtering for Micro-Diffusion Studies

THOMAS F. FISHER AND C. E. WEBER
*General Electric Company, Knolls Atomic Power Laboratory, Schenectady, New York**
(Received June 18, 1951)

A method which has been developed to allow uniform removal of micro-layers of metal from cylindrical specimens is based on sputtering of the metal from the cathode of a gas glow discharge tube. This technique is directly applicable to diffusion studies and offers many advantages over other more common methods.

INTRODUCTION

IF a metal specimen is made the cathode in a gas glow discharge tube, metal can be stripped from the surface in extremely thin, uniform layers. This method is being applied to diffusion studies by chemically analyzing the sputtered metal which is deposited on a collecting anode.

. Cathodic sputtering has many advantages over the more common methods of uniformly removing surface layers of metal. Machining generally requires that relatively thick layers be removed. Although chemical and electrochemical methods permit the removal of thin layers, extensive experimentation may be required to determine the proper solutions and concentrations, and further, the subsequent chemical analysis may be impeded by the presence of undesirable foreign ions. Sputtering is a generally applicable method and the sputtered metal can be collected on an inert or chemically compatible anode and dissolved in a minimum amount of solution. Sputtering can also be applied to the study of radioactive specimens where normal handling procedures are difficult or hazardous; it is easy to restrict spread of contaminating amounts of activity in a closed glass system during the sputtering operation, and the method is readily adaptable to remote operation.

A detailed discussion of the nature and theories of the mechanism of cathodic sputtering is available in the literature[1,2] and will not be presented in this article. However, a brief resumé may demonstrate how such a method can give a highly uniform removal of metal. In a steady state glow discharge, positively ionized gas atoms are accelerated to the cathode where they give up their energy. The surface atoms of the metal receive a portion of this energy and escape, diffusing to the walls of the system or to the anode[3–6]. The Crookes dark space, the region surrounding the cathode through which the incoming positive ions are accelerated, is a region of uniform electrical field where most of the voltage drop across the discharge tube occurs. In falling through this potential, the ions make a number of collisions so that their arrival at the cathode is neither mono-energetic nor uni-directional. The sputtering which occurs should be sub-microscopic in nature since the penetration of each incident ion is, at most, only a few atom layers. Due to the randomness of the events, the macroscopic dimension changes should be highly uniform.

EXPERIMENTAL PROCEDURE

The geometry of the sputtering system used in this investigation proved to be more critical than was first anticipated. Attempts to use planar geometry were not successful; cylindrical symmetry was necessary in order to achieve uniform metal removal. With flat specimens, the sputtering was not uniform across the specimen surface and proved to be dependent upon the shape and size of the glass system. When a large glass system was used, the removal of metal was principally on the edges of the specimen, while attack occurred in a small central region of the specimen when a small system was used.

With cylindrical geometry, the size of the system is apparently not critical, but the gas pressure, electrical potential, and cathode-anode distance are functionally related variables. It has been shown[7] that the product of the mass of sputtered materials, m (milligrams per ampere hour), the gas pressure, p (millimeters of mercury), and the distance between cathode and anode, D (centimeters), is a constant. In our work, the experimental apparatus and the sputtering conditions have been standardized. These conditions are given in Table I.

Under the above conditions, the cathode glow completely covered the cathode with the Crookes dark

TABLE I. Conditions for cathodic sputtering.

Sputtering gas	Krypton
Gas pressure	300 microns
Cathode to anode potential	1500 volts
Cathode to anode distance	0.5 inch
Specimen diameter	0.050–0.100 in.
Specimen length	0.50–0.75 in.
Current	1 milliampere

* The Knolls Atomic Power Laboratory is operated by the General Electric Company for the AEC. The work reported here was carried out under contract No. W-31-109 Eng-52.

[1] L. B. Loeb, *Fundamental Processes of Electrical Discharge in Gases* (John Wiley and Sons, Inc., New York, 1939).

[2] J. D. Cobine, *Gaseous Conductors, Theory and Engineering Applications* (McGraw-Hill Book Company, Inc., New York, 1941), first edition.

[3] A. V. Hippel, Ann. Physik **82**, 1043 (1936).

[4] K. H. Kingdon and I. Langmuir, Phys. Rev. **22**, 148 (1923).

[5] C. Starr, Phys. Rev. **56**, 216 (1939).

[6] C. H. Townes, Phys. Rev. **65**, 325 (1944).

[7] A. Güntherschulze, Z. Physik **36**, 563 (1926).

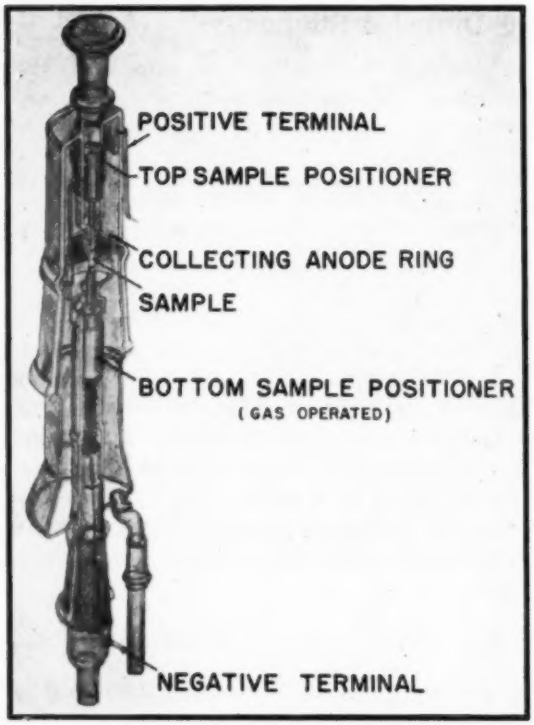


FIG. 1. Sputtering apparatus. (The cylindrical sample, which acted as the cathode, was centered vertically and radially in the cylindrical collecting anode. The supporting rods were cylindrical and had about the same diameter as the sample. The small diameter sample supports were attached to larger diameter pistons which operated vertically inside of glass cylinders.

The bottom piston was air operated for ease in loading the unit. With the top half of the sputterer removed, the lower piston could be forced down against the return spring by air pumped past the piston and into the connecting vacuum system. The piston retracted far enough into a glass collar to provide a space in the glass cylinder above the top of the lower supporting rod into which the sample was loaded. The sputterer top was then lowered into position, the top supporting cylinder in the down position fell directly on the glass cylinder containing the sample. When a vacuum seal was made and the gas pressure equalized on both sides of the lower piston, the return spring forced the piston upward and brought the sample into the sputtering position.)

space nearly filling the void between the cathode and the anode.

The sputterer used in these experiments (Fig. 1) was designed and constructed around a 45/50 standard taper ground glass joint.

During sputtering, most of the sputtered metal diffused to the collecting anode where it deposited. The inner surface of this anode ring was conveniently lined with a thin foil of metal which was easily removed after removal of each layer of metal, and the relative amount of each component was determined by chemical analysis. To date, the foil liner has been made of aluminum, which is easily dissolved and does not interfere with most chemical analyses. Any source of direct current was found to be suitable (partially rectified ac was satisfactory) with a protective resistance in series to prevent occurrence of excessively high currents during minor fluctuations in the glow discharge. The usual precautions to insure high purity krypton for sputtering were observed.

UNIFORMITY OF THE SPUTTERING PROCESS

Up to the present time, no prolonged attempt has been made to establish the limiting degree of uniformity attainable by cathodic sputtering. However, it is at least as good as the roundness of the best specimens available (i.e., ± 0.0001 inch on the diameter). Typical data are shown in Table II.

The samples were solid metal cylinders 0.50 in. in length, and the diameters were measured at random positions along the length of the sample. The minimum thickness that can be removed has yet to be determined, due to the eccentricity of the samples and resolving power limits on the optical measuring method used.

Preliminary runs showed that it is necessary to support the specimens at both ends to achieve uniform metal removal. A specimen originally 0.0740 in. O.D. was supported at one end and sputtered. The final diameters were 0.0680 in. near the supported end and 0.0625 in. near the unsupported end, a difference of 8 percent. This effect must have been due to the difference in electrical field around the exposed end of the sample resulting in a higher positive ion current in this region.

TABLE II. Uniformity of sputtering.

Sample	Diameter (before sputtering)	Diameter (after sputtering)
Stainless steel (S.A.E. Type 347)	0.0750 in.$\pm$0.0001 in.	0.0574 in.$\pm$0.0001 in.
Nickel on copper	0.0735 in.$\pm$0.0001 in.	0.0606 in.$\pm$0.0001 in.

SURFACE APPEARANCE AFTER SPUTTERING

A typical surface of a sputtered specimen is shown in Fig. 2. While it appears rough in the photograph, close inspection has shown the surface to be smooth. The depth of these irregularities appears to be less than 1×10^{-4} inches, but the manner in which light is reflected magnifies the surface.

A metallographic polishing action during sputtering has not been found. Apparently the incident ions are not affected by small electrical field differences resulting from the small surface irregularities, or at least not in a manner to cause polishing. The method has, however, been used to etch polished specimens such as zirconium, titanium, and beryllium in a method analogous to that described by McCutcheon.[9] Different grains and phases develop different appearances due to the sputtering action.

[9] D. M. McCutcheon and W. Pahl, Metals Progress **56**, 674 (1949).

SPUTTERING RATES

The factors which determine the rate of sputtering of the cathode in a gas glow discharge tube have been studied,[4–7] and perhaps the rate can be expressed most fundamentally as the atom yield (atoms sputtered per incident gas ion). However, for the purposes of this paper, an arbitrary rate expressed as mils of metal removed from the diameter of the specimen per milliampere-hour will be used, since theory alone does not permit calculation of atom yields from weight losses.

a. Specimen Temperature

Under any fixed set of conditions, sputtering rates increase rapidly with an increase in the specimen temperature. The application of sputtering to diffusion studies is experimentally restricted because the average specimen temperature must not be high enough for an appreciable amount of thermal diffusion to occur during sputtering. It should be noted that the average temperature of the specimen is mentioned rather than the high local temperature assumed to exist in the surface layer wherein the incident gas ions lose their energy. Preliminary experiments which were made to check the average temperature during sputtering by means of a thermocouple inserted into a hollow specimen gave temperatures of 225°C to 325°C with a current of one milliampere, under standard conditions. At this temperature, thermal diffusion is negligible for most metals.

TABLE III. Sputtering rates.

Specimen	Sputtering rate (mils/ma–hr)
Be	0.01 ±0.003
Mg	0.04 ±0.02
Al (S.A.E. 17S)	0.03 ±0.004
Cu	0.04 ···
Zr	0.05 ±0.01
Stainless steel (Type 347)	0.05 ±0.01
Ti	0.03 ±0.01
Ni[a]	0.049±0.001

[a] Closer control over gas purity and pressure in this series.

b. Experimental Sputtering Rates

To apply the sputtering technique to controlled removal of metal, the sputtering rate must be predetermined. A number of metals have been examined (Table III). With the exception of nickel, sputtering rates were measured with an accuracy of approximately 20 percent. The rate for nickel was obtained with extremely close control of the gas pressure and the gas purity, and is accurate to a few percent. However, any specific amount of metal can be removed using these rates as a basis for predicting sputtering times combined with intermittent determinations of the specimen diameter during sputtering. In operation, the amount of metal that can be removed at any one time is limited to two to three mils because the sputtered metal also deposits on the glass shield below the specimen and eventually shorts the system electrically. This shield is used to confine the sputtered metal, and if cleaned between runs can be used indefinitely.

FIG. 2. Surface appearance after sputtering.

The metals were of commercial purity in all cases. Beryllium metal, which sputtered more slowly than any of the other metals examined, was used for the sample supports in the sputtering apparatus and provided a longer life before replacement.

Pure metals sputter at different rates under similar conditions, which raises the question as to whether homogeneous metal alloys can be uniformly sputtered. For example, the selective removal of one component from an alloy surface may invalidate the method for diffusion studies. One can easily convince himself that no such effect can occur and this conclusion has been confirmed by experiments upon the two alloys, type 347 stainless steel, and 70–30 brass. The sputtered material from each of these alloys has been found to have the composition of the original material. The sputtering rates of the constituents of type 347 stainless steel (Fe, Ni, and Cr with a small amount of Cb) are probably about the same, but this is not at all true for copper and zinc, which make up the brass. In a heterogeneous system with large regions (compared to the gas ion penetrations) having dissimilar compositions, one region might, however, sputter faster than the other.

CONCLUSION

The general applicability of cathodic sputtering to qualitative removal of micro-quantities of metal has been demonstrated. Work is in progress to determine the limiting accuracy of the method and to apply it to thermal diffusion studies on metallic systems.

ACKNOWLEDGMENT

The authors wish to express their appreciation for the advice, counsel, and assistance given by Drs. J. E. Burke, J. P. Howe, and J. R. Low of the Knolls Atomic Power Laboratory of the General Electric Company.

Unsteady-State Separation Performance of Cascades. I

ELLIOTT W. MONTROLL AND GORDON F. NEWELL
Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland

(Received July 2, 1951)

This paper concerns the exact solution of the nonlinear differential equations which describe the time dependent behavior of multistage cascade separating processes for separation of two very similar molecular species. The Rayleigh separation law is postulated for each stage.

The main questions of interest in the nonstationary operation of cascades involve (a) their behavior while the concentration of the required component is being built up to the desired value; (b) the behavior of the cascade during the transition from one stationary mode of operation to another; and (c) the manner in which local concentration fluctuations are propagated through the cascade. We have derived analytical expressions that are suitable for the discussion of all of these points in square cascades. The cases of finite, half infinite, and infinite cascades are considered with and without product withdrawal.

A brief discussion of the linearization of a class of nonlinear second-order partial differential equations is given in Appendix I.

THE fractionation of mixtures of nearly identical components by multistage cascades is of considerable industrial importance. Several large scale multistage isotope separation plants which use the principles of thermal diffusion and molecular effusion have been built.[1] Distillation columns capable of separating isomers and other closely related components whose boiling points are almost the same and whose relative volatilities are close to unity are becoming more and more common. Solvent extraction processes for separating similar materials have been developed, and separation of similar gases by molecular diffusion is being investigated.[2]

The quantity of material in large cascades is frequently much greater than the normal daily output. Under such conditions the plant responds very slowly to changes in operating procedures or of equipment characteristics. The effect of the changes may not be reflected in the output for several weeks or even months. Compensating resettings of the operating characteristics might not restore the plant's normal output until several weeks later.

A significant fraction of the operating life of a large sluggish cascade is spent in the starting operation, in the transition from one set of operating conditions to another, and in recovering from transient disturbances (if there is a regular plant maintenance program, transient disturbances might occur frequently). Hence it is as important to develop optimum start-up and transition procedures and to be able to minimize transient losses in output as it is to set optimum steady-state operating conditions in the initial plant design.

In principle the unsteady-state behavior of a cascade can be deduced from the partial differential equations derived in Section 1. Unfortunately these equations are nonlinear, and in a typical cascade with stripper, a variable withdrawal rate, and sections with different characteristics, they are quite difficult to solve. Even when exact solutions can be obtained, they are sufficiently complicated so that they are not suited for approximate engineering calculations. It is therefore of interest to seek semi-quantitative methods of simply describing unsteady-state behavior of cascades as well as exact solutions of the time dependent separation equations.

This paper will be concerned with the determination of exact solutions of the nonlinear separation equations by making a transformation in the dependent variables of the type recently developed by H. Thomas,[6] E. Hopf,[8] and C. D. Majumdar.[9]

Since it is relatively easy to deduce the steady-state behavior of a cascade under almost any conditions, we shall, in the second paper of this series (by Dr. A. M. Squires and one of the authors E.W.M.), develop approximate formulas for the nonstationary behavior in terms of sets of related steady-state situations. Roughly speaking, we shall show that the steady state resulting from a nonstationary situation can be approached through a sequence of steady states corresponding to different production rates.

Since general discussions of cascade processes can be obtained in the literature, we shall give only a brief derivation of the fundamental equations which describe the variation in the concentration of one component of a binary mixture as a function of time and stage number. For detailed analyses of the considerations which must go into the selection of the cascade characteristics for a particular separation the reader is referred to the comprehensive reviews of R. Jones and W. Furry[3] (this review contains an extensive bibliography also),

[1] H. D. Smyth, *Atomic Energy for Military Purposes* (Princeton University Press, Princeton, New Jersey, 1945). See also J. F. Hogerton, Chem. Eng. **52**, 98 (1945) and P. C. Keith, Chem. Eng. **53**, 112 (1946).

[2] Cf. M. Benedict and A. Boas, Chem. Eng. Progress **47**, 111 (1951). Many general aspects of the engineering of cascade processes are discussed in these papers and in the article on Diffusion Separation Methods by M. Benedict in the *Encyclopedia of Chemical Technology* (Interscience Press, New York, 1950), Vol. V, p. 76.

[3] R. C. Jones and W. H. Furry, Revs. Modern Phys. **18**, 151 (1946).

K. Cohen,[4] and M. Benedict.[2] Although the Jones and Furry article is concerned mostly with separation by thermal diffusion, many of the results are applicable to other separation processes. Cohen's work contains the general theory of stationary operation of cascades. Benedict's report is given from more of an engineering point of view than either of the other two.

In this paper we shall consider the separation of a binary mixture whose components are represented by A and B. A is chosen to be the component of interest that is to be isolated from B. The part of the cascade where material enriched in A is to be withdrawn will be called the top; the part most depleted of A (in the stationary state), the bottom. It will be assumed that the unseparated mixture of A and B is fed into the bottom of the cascade. The authors hope to discuss the effect of a stripping section on the unsteady behavior of cascades at a later date.

1. DESCRIPTION OF CASCADE AND DERIVATION OF FUNDAMENTAL EQUATIONS

Let us consider a cascade of N separators (which we shall refer to as "stages") connected in the linear manner sketched in Fig. 1. A mixture of components A and B is fed into each stage and two streams, one enriched in A and one depleted of A are emitted from each stage. The enriched stream is fed into a "higher" stage and the depleted into a "lower" one. The following quantities characterize the operation of the nth stage at time T:

$W_1(n)$ = total rate of flow of material (A and B) in the nth enriched stream

$W_2(n)$ = total rate of flow of material (A and B) in the nth depleted stream

$y_n(T)$ = mol fraction of A in enriched stream ($1-y_n(T)$ = that of B)

$c_n(T)$ = mol fraction of A in depleted stream ($1-c_n(T)$ = that of B)

U_n = net flow of component A through dotted line which cuts cascade above the nth stage in Fig. 1.

H_n = capacity of nth stage.

If P is the rate at which material is withdrawn from the top of the plant we see that at all times

$$P=W_1(n)-W_2(n+1), \tag{1.1}$$

if the mixture of A and B in all concentrations is an incompressible fluid. The net flow U_n of component A at time t is

$$U_n=y_nW_1(n)-c_{n+1}W_2(n+1). \tag{1.2}$$

We shall postulate the law of separation of a given stage to be given by Rayleigh's formula[5]

$$\frac{y_n/(1-y_n)}{c_n/(1-c_n)}=e^{\psi}, \tag{1.3a}$$

where ψ is a constant independent of the concentrations y_n and c_n. ψ is called the over-all separation factor of the stage. It will be assumed that it is the same for every stage.

The over-all separation factor, ψ, is a very small number in cascades which require a large number of stages to perform the desired separation. The ratio $y_n/(1-y_n)$ is called the abundance ratio of A in the enriched stream of the nth stage; $c_n/(1-c_n)$, that in the depleted stream. As $y_n\to 0$, the abundance ratio approaches y_n.

When ψ is very small the following alternative form of (1.3a) is useful

$$\begin{aligned} y_n&=c_n(\exp\psi)/[1-c_n(1-\exp\psi)] \\ &=c_n+\psi c_n(1-c_n)+O(\psi^2). \end{aligned} \tag{1.3b}$$

We shall assume in the remainder of this paper that ψ is so small that terms of order ψ^2 can be neglected in (1.3b).

The rate of accumulation of A in the nth stage is the difference between the rate of inflow and that of emission of A from that stage. That is,

$$\begin{aligned} H_n\partial c_n/\partial T&=y_{n-1}W_1(n-1)+c_{n+1}W_2(n+1) \\ &\qquad -c_nW_2(n)-y_nW_1(n) \\ &=\Delta\{c_nW_2(n)-y_{n-1}W_1(n-1)\}, \end{aligned} \tag{1.4}$$

where Δ represents the difference operator with the property $\Delta a_n=a_{n+1}-a_n$. By substituting (1.1) and (1.3b) into (1.4) we obtain an equation which involves only H_n, c_n, $W_1(n)$ and the constants P and ψ. Then

$$\begin{aligned} &H_n\partial c_n/\partial T \\ &\quad=\Delta\{W_1(n-1)[\Delta c_{n-1}-\psi c_{n-1}(1-c_{n-1})]-Pc_n\}. \end{aligned} \tag{1.5}$$

One can find the concentration of A as a function of n and T by solving (1.5) with the appropriate boundary conditions. Now c_n is a very slowly varying function of n (as is $W_1(n)$) in cascades with small over-all separation per stage. Under these conditions the difference operator can be replaced by a differential operator and (1.5) becomes the differential equation

$$H\partial c/\partial T=\partial/\partial n\{W_1[\partial c/\partial n-\psi c(1-c)]-Pc\}. \tag{1.6}$$

This is our fundamental equation.

The boundary condition at the top of a cascade is a statement of the observation that $y_N\equiv c_{N+1}$ in Fig. 1. That is, the enriched material at the top stage is partly withdrawn and partly recycled into the top stage. Then

[4] K. Cohen, *Science and Engineering of Nuclear Power* (Addison-Wesley Press Inc., Cambridge, Mass., 1949), Vol. 2, Chapter 2, p. 19. See also R. Peierls, *Science News* (Penguin Books, London, 1947), Vol. 2, p. 41.

[5] Lord Rayleigh, Phil. Mag. 42, 493 (1896) or *Collected Works* (Cambridge University Press, London), Vol. 4, p. 261.

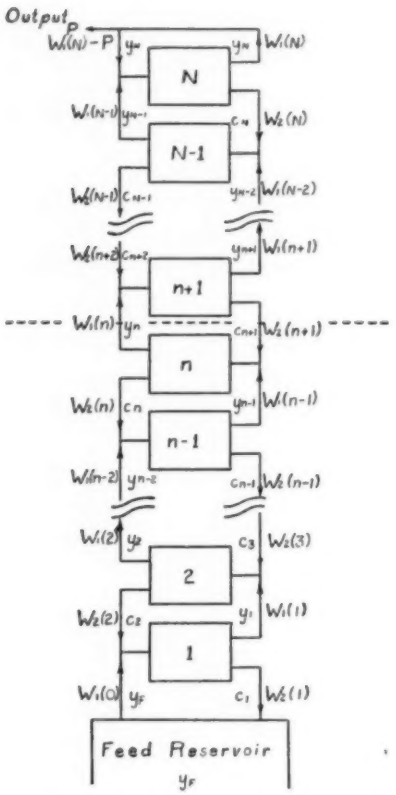


FIG. 1. Flow diagram of cascade of separators. W_2's represent flow rates of depleted streams, W_1's that of enriched streams, y_n's and c_n's the mol fraction of component of interest in enriched stream.

the Rayleigh separation formula (1.3b) is equivalent to

$$c_{N+1}-c_N=\Delta c_N=\psi c_N(1-c_N). \tag{1.7a}$$

Again we replace the difference operator by the differential operator to obtain

$$\partial c/\partial n=\psi c(1-c) \quad \text{when} \quad n=N \quad \text{and} \quad T>0. \tag{1.7b}$$

We shall use the boundary condition at the 0th stage (bottom of cascade) which corresponds to a cascade fed from an infinite reservoir of material whose concentration of A is y_F. Since for small ψ the concentration of A in the depleted stream is almost y_F, we can write

$$c(0)=y_F \quad \text{for} \quad T>0. \tag{1.8}$$

2. STATIONARY OPERATION OF CASCADES

During stationary cascade operation $\partial c/\partial T=0$. Hence

$$W_1(n)\{(\partial c/\partial n)-\psi c(1-c)\}=Pc+\text{constant}. \tag{2.1}$$

The physical significance of the constant can be determined from the boundary condition (1.7). We find

$$W_1(n)\{(\partial c/\partial n)-\psi c(1-c)\}=P(c-y_p), \tag{2.2}$$

where y_p is the mol fraction of A in the product withdrawn from the top of the cascade.

In a cascade from which no product is withdrawn ($P=0$) and in which $W_1(n)\neq 0$ for $0\leq n\leq N$, (2.2) reduces to

$$\partial c/\partial n=\psi c(1-c), \tag{2.3}$$

with the boundary condition $c(0)=y_F=$ feed concentration. Then at the nth stage

$$c/(1-c)=[y_F/(1-y_F)]\exp\psi n,$$

so that

$$c(n)=y_F e^{n\psi}/[1+y_F(e^{n\psi}-1)]. \tag{2.4}$$

It is sometimes convenient to number the stages from the top, calling $m=0$ the top stage, $m=-1$, the stage next to the top, etc. Then $m=n-N$ and

$$c(m)=y_F e^{m\psi}/[e^{-N\psi}(1-y_F)+y_F e^{m\psi}]. \tag{2.5}$$

As $N\to\infty$ (i.e., in a semi-infinite plant)

$$c(m)=1 \quad \text{for} \quad m>-\infty. \tag{2.6}$$

Equation (2.2) can also be integrated easily when $W_1(n)=\text{constant}=L$. A cascade with this property is called a *square cascade*. The remaining sections of this paper will be concerned mainly with the nonequilibrium behavior of square cascades.

The combinations

$$\kappa=P/L\psi, \quad \alpha_1=\tfrac{1}{2}(1-\kappa), \quad \text{and} \quad \alpha_2=\tfrac{1}{2}(1+\kappa), \tag{2.7a}$$

will appear frequently. It will also be convenient to use the quantities

$$u(n)=c(n)-\alpha_2, \tag{2.7b}$$

and

$$\gamma=\{\alpha_2^2-\kappa y_p\}^{\frac{1}{2}}. \tag{2.8}$$

Since $0\leq y_p\leq 1$, we have

$$\alpha_1\leq\gamma\leq\alpha_2. \tag{2.9}$$

In most cases of interest $\kappa\ll 1$. The differential Eq. (2.2) is equivalent to

$$\partial u/\partial(n\psi)=\gamma^2-u^2, \tag{2.10}$$

and the boundary condition (1.7b) at $n=N$ becomes

$$\partial n/\partial(n\psi)=-(u-\alpha_1)(u+\alpha_2) \quad \text{at} \quad n=N. \tag{2.10a}$$

The solution of (2.10) is

$$u(n)=\gamma\tanh(n\psi\gamma+\delta_0)=c(n)-\alpha_2, \tag{2.11}$$

where

$$\delta_0=\tanh^{-1}(u_0/\gamma), \tag{2.12a}$$

with

$$u_0=y_F-\alpha_2. \tag{2.12b}$$

One can see from the Rayleigh separation formula (1.3b) that it is desirable to choose $W_1(n)$ to be large when n is small and vice versa. When c_n and y_n are both small compared with 1, as they would be near the feed stages, $y_n/c_n\simeq(1+\psi)$. On the other hand, when c_n and y_n are close to unity, as they would be near the top of a long cascade, $y_n/c_n\simeq 1+\psi(1-c_n)$. Hence, there is a much greater relative separation per stage at the

bottom stages than there is at the top. On this basis one should recycle with a greater frequency at the lower stages than in the upper ones. If a given amount of power is available for cascade operation, the larger fraction of it should be directed toward recycling in the lower stages. This is equivalent to saying that $W_1(n)$ should be largest at stages in the neighborhood of the feed. Also, when highly separated material is valuable, it is desirable to make the hold up H_n small for large n.

A rational choice of the function $W_1(n)$ can be made by designing the cascade in such a way that no mixing of materials of different concentrations occurs at the feed point of the nth stage where the stream depleted of A from the $(n+1)$-st stage mixes with the enriched fraction from the $(n-1)$-st. A cascade with this property at every stage is called an ideal cascade. It can be shown[4] that in this case $W_1(n)$ is a monotone decreasing function of n. For the sake of over-all economy and operating flexibility it is desirable to make cascades neither ideal nor square, but rather as a hybrid of the two. In order to make $W_1(n)$ a continuous decreasing function of n when N is large (say $O(10^3)$), a tremendous number of equipment sizes or control settings, or both, is required. Since $W_1(n)$ in an ideal cascade depends on the withdrawal rate, P, all flow rates must be changed to preserve ideality when P is changed. The economical hybrid cascade consists of several square cascades in series. The theory of such a system is complicated because a set of connecting conditions exists at the points where one square cascade ends and another begins. The theory would be analogous to that of heat flow through laminated material in which each layer had a different thermal conductivity.

3. LINEARIZATION OF FUNDAMENTAL EQUATION

In a square cascade with $W_1(n)=L$ and $H_n=H$ for $0\leq n\leq N$, the fundamental differential equation (1.6) is

$$2\tau\partial c/\partial T=\partial/\partial n\{\partial c/\partial n-\psi c(1-c)+Pc/L\}, \quad (3.1)$$

where $\tau=H/2L=$"hold up time"$\simeq$average time spent by a molecule in a given stage during one passage through the stage. The boundary conditions in a cascade over an infinite reservoir are

$$c(0, T)=y_F \quad \text{for all} \quad T>0, \quad (3.2)$$

$$\partial c/\partial n=\psi c(1-c) \quad \text{for} \quad n=N,\ T>0. \quad (3.3)$$

We shall postulate the initial concentration distribution to be

$$c(n, 0)\equiv c_0(n) \quad \text{at} \quad T=0. \quad (3.4)$$

Although (3.1) is nonlinear, it can easily be linearized by a transformation of the independent variable. Apparently this technique was first used in a separation problem by H. Thomas[6] in his paper on ion exchange in flowing systems: By making the transformation to a new variable $\Phi(x, y)$

$$Cf=\log\{e^{-(Bx+Ay)}\Phi(x, y)\},$$

he solves the nonlinear equation

$$f_{xy}+Af_x+Bf_x+Cf_xf_y=0,$$

where $f_x=\partial f/\partial x$, etc., and A, B, and C are constants. The equation which describes the Burgers[7] one dimensional model of turbulence $u_t=(\mu u_x-\frac{1}{2}u^2)_x$ (with μ constant) has been linearized by E. Hopf.[8] C. D. Majumdar[9] used essentially the same transformation to solve the problem in thermal diffusion separation which is equivalent to the determination of concentration as a function of n and T in a square cascade filled with material of mol fraction y_F of A. He considers only the case of no withdrawal from the top of the thermal diffusion column. It is interesting that all the workers mentioned above independently used essentially the same methods to solve three different problems. We will make a few remarks about the general linearization schemes in Appendix I.

Equation (3.1) can be transformed into a more convenient form by introducing the variables defined in (2.7) and the two new variables

$$x=n\psi \quad \text{and} \quad t=T\psi^2/2\tau. \quad (3.5)$$

Then (3.1) becomes

$$u_t=(u_x+u^2)_x. \quad (3.6)$$

Proceeding in the manner of Hopf and Majumdar we introduce a new dependent variable which is defined by

$$\varphi(x, t)=\exp\int u dx. \quad (3.7a)$$

The inverse of this substitution is

$$u=(\log\varphi)_x. \quad (3.7b)$$

Then (3.6) becomes

$$(\log\varphi)_{tx}=(\varphi_{xx}/\varphi)_x, \quad (3.8)$$

which, upon integration with respect to x, yields

$$\varphi_t=\varphi_{xx}+c(t)\varphi,$$

where $c(t)$ is an arbitrary function of t. This function could be eliminated by introducing a new variable $\varphi^*=\varphi\exp\{-\int c(t)dt\}$, which would satisfy $\varphi_t^*=\varphi_{xx}^*$. However, since $(\log\varphi^*)_x=(\log\varphi)_x$, Eq. (3.7b) shows that u and hence the concentration, c, is independent of $c(t)$. We can therefore choose $c(t)\equiv 0$ for all t. Then our fundamental differential equation (3.1) is reduced to the heat conduction equation

$$\varphi_t=\varphi_{xx}. \quad (3.9)$$

[6] H. Thomas, J. Am. Chem. Soc. **66**, 1664 (1944).

[7] J. M. Burgers, Proc. Acad. Sci. Amsterdam **43**, 2 (1940); **53**, 247 (1950).

[8] E. Hopf, Comm. Pure Ap. Math. **3**, 201 (1950). Apparently the required transformation has also been known to C. D. Cole and V. Bargmann (see footnote on p. 252 of Burgers' 1950 paper).

[9] C. D. Majumdar, Phys. Rev. **81**, 844 (1951). Majumdar also linearized $2\nabla^2\varphi=\varphi_x^2+\varphi_y^2+\varphi_z^2$ which occurs in a problem in theory of general relativity Phys. Rev. **72**, 390 (1947).

Fortunately (3.7) also linearizes the boundary conditions (3.3) to

$$\varphi_{xx}+\kappa\varphi_x-\alpha_1\alpha_2\varphi=0 \quad \text{at} \quad x=N\psi=l, \quad t>0, \tag{3.10a}$$

$$\varphi_x=(\gamma_F-\alpha_2)\varphi \quad \text{at} \quad x=0, \quad t>0, \tag{3.10b}$$

$$\varphi(x,0)=\exp\left\{\int_0^x[c_0(z/\psi)-\alpha_2]dz\right\}, \quad t=0, \quad x>0. \tag{3.10c}$$

Since these boundary conditions contain second derivatives, they are not of the usual Sturm-Liouville type. We shall make transformations which put them in a more classical form.

4. INFINITE CASCADE

The manner in which a concentration fluctuation in a region of a cascade far from either end is propagated through the cascade can be discussed in terms of an infinite cascade. Then one must solve (3.9) with the initial condition (3.10c). The range of x is $-\infty<x<\infty$.

Before finding the exact solution of the nonlinear separation equation let us consider the case of an infinite cascade whose mole fraction of A at every stage is initially $\ll 1$. Under this condition we can use the linearized form of (3.1) which is obtained by neglecting c^2 as compared with c:

$$c_t=\{c_x-c(1+\kappa)\}_x. \tag{4.1}$$

Since this equation is linear, we can interpret c as either the concentration of A or as the deviation of the concentration of A from any stationary solution of (4.1). Equation (4.1) can be transformed into the heat conduction equation in the new variable f defined by

$$f=c\exp\{\alpha_2^2t-\alpha_2x\}.$$

By employing the method of Fourier transforms it is easy to show[10] that in an infinite cascade

$$c(x,t)=\tfrac{1}{2}(\pi t)^{-\frac{1}{2}}e^{-\alpha_2^2t}\times\int_{-\infty}^{\infty}c_0(z)e^{-\alpha_2(z-x)}e^{-(1/4t)(z-x)^2}dz, \tag{4.2}$$

where $c_0(x)$ is the initial concentration distribution of A at time $t=0$, or the deviation of the initial concentration distribution from a stationary state. Equation (4.2) is equivalent to

$$c(x,t)=\tfrac{1}{2}(\pi t)^{-\frac{1}{2}}\int_{-\infty}^{\infty}c_0([x-t(1+\kappa)]+v)e^{-(1/4t)v^2}dv.$$

To see how a packet of A is propagated through the cascade we choose $c_0(x)$ to be proportional to a δ-function. $c_0(x)=\epsilon\delta(x)=0$ unless $x=0$, and $\int\delta(x)dx=1$ (ϵ is a small constant). Then

$$c(x,t)=\tfrac{1}{2}\epsilon(\pi t)^{-\frac{1}{2}}\exp-\{(1/4t)[x-t(1+\kappa)^2]\}.$$

Initially the peak of the distribution is at $x=0$. At time t it is at the point which makes the argument of the exponential function vanish, $x=t(1+\kappa)$. Hence, in our original time and stage number scale the rate at which the peak is propagated up the cascade is

$$dn/dT=(\psi/2\tau)[1+(P/F\psi)].$$

A measure of the spread of the packet with time is given by the standard deviation of $c(x,t)$ from the mean. This is $\sigma=\psi(T/\tau)^{\frac{1}{2}}$, so that the rate of spread of the packet is $d\sigma/dT=\tfrac{1}{2}\psi(\tau T)^{-\frac{1}{2}}$. Similar results have been obtained by A. Squires and H. Callen for the mode of propagation of contaminant through a cascade.

The exact variation of $c(x,t)$ in an infinite cascade is determined by solving $\varphi_t=\varphi_{xx}$ with the initial distribution

$$\varphi(x,0)=\exp\int_0^x[c_0(z/\psi)-\alpha_2]dz,$$

by the usual Fourier transform method[10] to obtain

$$\varphi(x,0)=\tfrac{1}{2}(\pi t^{-\frac{1}{2}})\int_{-\infty}^{\infty}\exp\left\{\int_0^u[c(z/\psi)-\alpha_2]dz-\tfrac{1}{4}(u-x)^2/t\right\}du. \tag{4.3}$$

The concentration of A as a function of x and t is then

$$c(x,t)=\alpha_2+[\log\varphi(x,t)]_x=\alpha_2-\frac{1}{2t}\frac{\int_{-\infty}^{\infty}(x-u)\exp\left\{\int_0^u[c(z/\psi)-\alpha_2]dz-\tfrac{1}{4}(x-u)^2/t\right\}du}{\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{4t}(x-u)^2+\int_0^u[c_0(z/\psi)-\alpha_2]dz\right\}dn}. \tag{4.4}$$

Transient local fluctuations in concentration can be followed through the application of this formula.[11]

[10] Titchmarsh, *Theory of Fourier Integrals* (Oxford University Press, London, 1937), p. 281.

[11] An analogous result has been obtained independently by W. H. Stockmayer in a problem concerning the theory of moving concentration boundaries in electrolytic cells. Such processes are also described by an equation similar to (3.6). This work is to be published in the Annals of the New York Academy of Science.

5. SEMI-INFINITE CASCADE

Suppose we consider a finite cascade that is initially uniformly filled with feed material whose mole fraction of A is y_F. The separation equation indicates that if $c=$constant for all x, then $\partial c/\partial T=0$ and c does not change with time. How then can separation start in this cascade if its feed stage is connected to an infinite reservoir whose concentration of A is y_F? From the boundary condition (3.3), $\partial c/\partial n=\psi c(1-c)$ at the top of the cascade, we see that at the top stage $\partial c/\partial n\neq 0$ except at $t=0$. Hence, separation starts at the top and works its way down the cascade. If one wishes to determine the manner in which separation of feed material starts, he needs only to determine the behavior of a semi-infinite cascade. This simplifies calculations considerably because the effect of the boundary condition at the feed stage can be ignored. The separating performance of a finite cascade can be approximated by that of a semi-infinite one until the variation of the concentration gradient from zero or its stationary value has been propagated to the feed stage. If $T\ll 2\tau N/\psi$ this condition is satisfied.

We number the stages of a semi-infinite cascade in such a way that $n=0$ (or $x=0$) corresponds to the top stage, $n=-1$ to that next to the top, etc. The boundary condition at the top is then

$$\varphi_{xx}+\kappa\varphi_x=\tfrac{1}{4}(1-\kappa^2)\varphi=\alpha_1\alpha_2\varphi \quad \text{at} \quad x=0. \tag{5.1a}$$

We let $\varphi_0(x)$ correspond to the initial concentration distribution. The behavior of the cascade is then to be deduced from the "heat conduction equation"

$$\varphi_{xx}=\varphi_t, \tag{5.1b}$$

with the initial condition,

$$\varphi(x,0)=\varphi_0(x), \tag{5.1c}$$

and the boundary condition (5.1a).

To find the complete solution of (5.1) we introduce the new function $v(x,t)$ defined by

$$v(x,t)=\varphi_{xx}+\kappa\varphi_x-\tfrac{1}{4}(1-\kappa^2)\varphi. \tag{5.2}$$

Then $v(x,t)$ is the solution of the classical problem[12]

$$v_{xx}=v_t, \tag{5.3}$$

$$v(0,t)=0, \tag{5.4}$$

$$v(x,0)=v_0(x)=\varphi_0''+\kappa\varphi_0'-\tfrac{1}{4}(1-\kappa^2)\varphi_0, \tag{5.5}$$

which corresponds to the flow of heat in a semi-infinite rod with one end kept at zero temperature and with the initial temperature distribution $v(x,0)=v_0(x)$.

[12] Carslaw, *Conduction of Heat in Solids* (Dover Publications, Inc., New York, 1945), p. 33.

The solution of (5.3) is[12]

$$v(x,t)=(1/2(\pi t)^{\frac{1}{2}})\int_{-\infty}^{0} v_0(x')\{\exp[-(x-x')^2/4t]-\exp[-(x+x')^2/4t]\}dx'$$

$$v(z,t)=(1/2(\pi t)^{\frac{1}{2}})\int_{-\infty}^{0}[\varphi_0''(x')+\kappa\varphi_0'(x')-\tfrac{1}{4}(1-\kappa^2)\varphi_0(x')]\{\exp[-(z-x')^2/4t]-\exp[-(z+x')^2/4t]\}dx'. \tag{5.6}$$

Given the analytical expression for $v(x,t)$ we can determine $\varphi(x,t)$ by solving the nonhomogeneous equation (5.2) which yields

$$\varphi(x,t)=A(t)e^{\alpha_1 x}+B(t)e^{-\alpha_1 x}+\int_0^x v(z,t)[e^{\alpha_1(x-z)}-e^{-\alpha_2(x-z)}]dz, \tag{5.7}$$

where

$$\alpha_1=\tfrac{1}{2}(1-\kappa) \quad \text{and} \quad \alpha_2=\tfrac{1}{2}(1+\kappa). \tag{5.8}$$

The arbitrary functions $A(t)$ and $B(t)$ are to be determined so that (5.7) actually satisfies (5.1a–5.1c). The following three identities will be useful in this determination:

$$v_x(0,t)=\tfrac{1}{2}(\pi t)^{-\frac{1}{2}}\int_{-\infty}^{0} v_0(y)(y/t)\exp(-y^2/4t)dy, \tag{5.9}$$

$$\int_0^x e^{\alpha_1(x-z)}(v_{zz}-v\alpha_1^2)dz=v_x(x,t)-v_x(0,t)e^{x\alpha_1}+\alpha_1 v(x,t), \tag{5.10a}$$

$$\int_0^x e^{-\alpha_2(x-z)}(v_{zz}-v\alpha_2^2)dz=v_x(x,t)-v_x(0,t)e^{-\alpha_2 x}-\alpha_2 v(x,t). \tag{5.10b}$$

Equations (5.9) follows directly from (5.6) whereas (5.10a) and (5.10b) are obtained by an integration by parts (with the use of $v(0,t)\equiv 0$).

Since $v_t=v_{zz}$, one obtains by differentiation of (5.7) and application of (5.9–5.10).

$$A'(t)-\alpha_1^2 A(t)=\tfrac{1}{2}(\pi t)^{-\frac{1}{2}}\int_{-\infty}^{0} v_0(z)(z/t)\exp(-z^2/4t)dz, \tag{5.11a}$$

$$B'(t)-\alpha_2^2 B(t)=-\tfrac{1}{2}(\pi t)^{-\frac{1}{2}}\int_{-\infty}^{0} v_0(z)(z/t)\exp(-z^2/4t)dz. \tag{5.11b}$$

The solutions of these first-order linear equations are

$$A(t)=\exp(t\alpha_1^2)\left\{A(0)-\pi^{-\frac{1}{2}}\int_0^t \tau^{-\frac{1}{2}}\exp(-\tau\alpha_1^2)d\tau \times\left[v_0(0)-\int_{-\infty}^0 v_0'(z)\exp(-z^2/4\tau)dz\right]\right\}, \quad (5.12a)$$

$$B(t)=\exp(t\alpha_2^2)\left\{B(0)+\pi^{-\frac{1}{2}}\int_0^t \tau^{-\frac{1}{2}}\exp(-\tau\alpha_2^2)d\tau \times\left[v_0(0)-\int_{-\infty}^0 v_0'(z)\exp(-z^2/4\tau)dz\right]\right\}. \quad (5.12b)$$

The constants $A(0)$ and $B(0)$ are found by substituting (5.12a), (5.12b), and (5.5) into (5.7) when $t=0$. Then (after several integrations by parts)

$$\varphi_0(x)=e^{x\alpha_1}[A(0)-\varphi_0'(0)-\alpha_2\varphi_1(0)] +e^{-x\alpha_2}[B(0)+\varphi_0'(0)-\alpha_1\varphi_0(0)]+\varphi_0(x).$$

This equation can be satisfied for all x only if

$$A(0)=\varphi_0'(0)+\alpha_2\varphi_0(0), \quad (5.13a)$$

$$B(0)=-\varphi_0'(0)+\alpha_1\varphi_0(0). \quad (5.13b)$$

The complete solution of the main boundary value problem (5.1) is (5.7) with $v(z, t)$ given by (5.6), $A(t)$ and $B(t)$ by (5.12), and $A(0)$ and $B(0)$ by (5.13). The mol fraction of A, $c(x, t)$ is the logarithmic derivative of $\varphi(x, t)$.

As a special application of these results let us consider the early time behavior of a cascade initially filled with feed of concentration y_F and with no withdrawal from the top. Then $c_0=y_F$, $u_0=y_F-\frac{1}{2}=-\sigma$ and

$$\varphi(x, 0)=\varphi_0(x)=\exp\int_0^x (y_F-\tfrac{1}{2})dx=\exp-\sigma x. \quad (5.14)$$

Clearly when $\kappa=0$,

$$v_0(x)=\varphi_0''-\tfrac{1}{4}\varphi_0(x)=(\sigma^2-\tfrac{1}{4})\exp(-\sigma x),$$

so that

$$v(z, t)=(\sigma^2-\tfrac{1}{4})/2(\pi t)^{\frac{1}{2}} \times\int_{-\infty}^0 \exp(-\sigma x')\{\exp[-(x'-z)^2/4t] -\exp[-(x'+z)^2/4t]\}dx', \quad (5.15)$$

$$\alpha_1=\alpha_2=\tfrac{1}{2};\quad A(0)=(\tfrac{1}{2}-\sigma)=y_F;\quad B(0)=(\tfrac{1}{2}+\sigma)=1-y_F. \quad (5.16)$$

Then (5.11) implies that

$$A(t)e^{x\alpha_1}+B(t)e^{-x\alpha_2}=e^{\frac{1}{2}x+\frac{1}{4}t}-2\pi^{-\frac{1}{2}}\sinh\tfrac{1}{2}x \times\left\{e^{\frac{1}{4}t}\int_{\frac{1}{2}(t)^{\frac{1}{2}}}^\infty e^{-u^2}du+2\sigma e^{t\sigma^2}\int_{-\infty}^{\sigma t^{\frac{1}{2}}}\exp(-u^2)du\right\}. \quad (5.17a)$$

We also have

$$\int_0^x v(x, t)[e^{\alpha_1(x-z)}-e^{-\alpha_2(x-z)}]dz=(1/2(\pi t)^{\frac{1}{2}}) \times\left\{4\sigma\sinh\tfrac{1}{2}x\int_{-\infty}^0 e^{-u\sigma}\exp(-u^2/4t)du +e^{-\sigma x}\int_{-\infty}^{-x} e^{-u\sigma}\exp(-u^2/4t)du -e^{+\sigma x}\int_{-\infty}^{x} e^{-u\sigma}\exp(-u^2/4t)du +e^{-\frac{1}{2}x}\int_0^x e^{\frac{1}{2}u}\exp(-u^2/4t)du +e^{\frac{1}{2}x}\int_0^x e^{-\frac{1}{2}u}\exp(-u^2/4t)du\right\}. \quad (5.17b)$$

Hence (after expressing the terms in (5.17b) as integrals of Gaussian functions)

$$\varphi(x, t)=e^{\frac{1}{2}x}e^{\frac{1}{4}t} +\pi^{-\frac{1}{2}}e^{\frac{1}{4}t}\left(e^{-\frac{1}{2}x}\int_{\frac{1}{2}t^{\frac{1}{2}}-\frac{1}{2}xt^{-\frac{1}{2}}}^\infty \exp(-u^2)du -e^{\frac{1}{2}x}\int_{\frac{1}{2}t^{\frac{1}{2}}+\frac{1}{2}xt^{-\frac{1}{2}}}^\infty \exp(-u^2)du\right) +\pi^{-\frac{1}{2}}e^{t\sigma^2}\left(e^{-\sigma x}\int_{\frac{1}{2}xt^{-\frac{1}{2}}-\sigma t^{\frac{1}{2}}}^\infty \exp(-u^2)du -e^{\sigma x}\int_{-\frac{1}{2}xt^{-\frac{1}{2}}-\sigma t^{\frac{1}{2}}}^\infty \exp(-u^2)du\right). \quad (5.18)$$

Equation (5.18) can be simplified by introducing the function

$$\Theta(x)=\pi^{-\frac{1}{2}}e^{x^2}\int_x^\infty \exp(-u^2)du =\tfrac{1}{2}\exp(x^2)(1-erfcx)=\tfrac{1}{2}\exp(x^2)erfx, \quad (5.19)$$

where

$$erfx=2\pi^{-\frac{1}{2}}\int_0^x \exp(-u^2)du,$$

and $erfcx$ is the so-called complementary error function $(1-erfx)$.

In terms of $\Theta(x)$ we have

$$\varphi(x, t)=\exp[\tfrac{1}{2}x+\tfrac{1}{4}t] +e^{-x^2/2t}\{\Theta(\tfrac{1}{2}t^{\frac{1}{2}}-\tfrac{1}{2}xt^{-\frac{1}{2}})-\Theta(\tfrac{1}{2}t^{\frac{1}{2}}+\tfrac{1}{2}xt^{-\frac{1}{2}}) +\Theta(\tfrac{1}{2}xt^{-\frac{1}{2}}-\sigma t^{\frac{1}{2}})-\Theta(-\tfrac{1}{2}xt^{-\frac{1}{2}}-\sigma t^{\frac{1}{2}})\}. \quad (5.20)$$

To determine the mol fraction of A as a function of x and t we use the relations (2.7) and (3.7b). Then for $x \leq 0$,

$$c(x,t)=1-\frac{(1-y_F)\Theta(\frac{1}{2}xt^{-\frac{1}{2}}-\sigma t^{\frac{1}{2}})-y_F\Theta(-\frac{1}{2}xt^{-\frac{1}{2}}-\sigma t^{\frac{1}{2}})+\Theta(\frac{1}{2}t^{\frac{1}{2}}-\frac{1}{2}xt^{-\frac{1}{2}})}{\Theta(\frac{1}{2}xt^{-\frac{1}{2}}-\sigma t^{\frac{1}{2}})-\Theta(-\frac{1}{2}xt^{-\frac{1}{2}}-\sigma t^{\frac{1}{2}})+\Theta(-\frac{1}{2}xt^{-\frac{1}{2}}+\frac{1}{2}t^{\frac{1}{2}})-\Theta(\frac{1}{2}xt^{-\frac{1}{2}}+\frac{1}{2}t^{\frac{1}{2}})+\exp(\frac{1}{2}xt^{-\frac{1}{2}}+\frac{1}{2}t^{\frac{1}{2}})^2}. \quad (5.21)$$

The mol fraction of A at the top stage ($x=0$) at time t has an especially simple form

$$c(0,t)=\tfrac{1}{2}+\tfrac{1}{2}erf(\tfrac{1}{2}t^{\frac{1}{2}})-\sigma e^{-(\frac{1}{4}-\sigma^2)t}[1+erf(\sigma t^{\frac{1}{2}})]. \quad (5.22)$$

6. FINITE CASCADE OVER INFINITE RESERVOIR

In this section we consider the variation of the mol fraction of A as a function of time and stage number in a finite square cascade fed from an infinite reservoir. For the general problem of the variation from an arbitrary initial distribution in a cascade with withdrawal rate P (corresponding to $\kappa=P/L\psi$) the fundamental equations are (after linearization)

$$\varphi_t=\varphi_{xx}, \quad (3.9)$$

$$\varphi_{xx}+\kappa\varphi_x-\alpha_1\alpha_2\varphi=0 \text{ at } x=N\psi=l \text{ for } t>0, \quad (3.10a)$$

$$\varphi_x=(y_F-\alpha_2)\varphi \text{ at } x=0,\ t>0, \quad (3.10b)$$

$$\varphi(x,0)=\exp\int_0^x[c_0(x/\psi)-\alpha_2]dx \text{ at } t=0,\ x>0. \quad (3.10c)$$

Unfortunately, if one attempts to separate the variables in the usual way, the linear ordinary differential equation in x with the boundary conditions (3.10a) and (3.10b) is not self-adjoint. This means that the characteristic functions of the appropriate linear operator do not form an orthogonal set. Hence, as the problem now stands, one has difficulty in fitting the initial conditions (3.10c).

In order to reduce 3.9–3.10 to a more classical self-adjoint problem[13] we follow the philosophy of Section 5 and introduce a new variable $\omega(x,t)$ which is defined in such a way that (3.10a) and (3.10b) are transformed to the more classical boundary conditions

$$\omega(0,t)=0, \quad (6.1a)$$

$$\omega_x+\delta\omega=0 \quad \text{for} \quad t>0,\ x=l. \quad (6.1b)$$

Here δ is a constant which is to be chosen later.

Let

$$\omega(x,t)=\varphi_x-f(x)\varphi(x,t), \quad (6.2)$$

where $f(x)$ is an arbitrary function of x. Then from (3.9) $\omega(x,t)$ satisfies the differential equation

$$\omega_{xx}=\omega_t-2\omega f_x-[f_{xx}+2ff_x]\varphi. \quad (6.3)$$

This is simplified by choosing $f(x)$ to be a solution of the ordinary differential equation

$$f_{xx}+2ff_x=0 \quad \text{or} \quad \partial/\partial x[f_x+f^2]=0, \quad (6.4)$$

for otherwise (6.3) would be an integro-differential equation in $\omega(x,t)$.

The boundary condition (3.10b) is equivalent to (6.1a) if

$$f(0)=(y_F-\alpha_2). \quad (6.5)$$

The boundary condition (3.10a) is equivalent to

$$\omega_x+(f+\kappa)\omega+\varphi[f_x+f(f+\kappa)-\alpha_1\alpha_2]=0. \quad (6.6)$$

Hence, if

$$\delta=f(l)+\kappa, \quad (6.7)$$

and

$$f_x+f(f+\kappa)-\alpha_1\alpha_2=0 \quad \text{for} \quad x=l, \quad (6.8)$$

then $\omega(x,t)$ satisfies our desired condition (6.1b). Now it is to be noticed that the differential equation (6.4) for $f(x)$, and the boundary conditions (6.5) and (6.8) are exactly the same as those satisfied by the stationary solution of the fundamental equation (2.10), (2.12), and (2.10a). Hence from (2.11)

$$f(x)=\gamma\tanh(x\gamma+\delta_0), \quad (6.9)$$

where δ_0 is given by (2.12) and γ by (2.8).

The differential equation (6.3) is then

$$\omega_t=\omega_{xx}+2\gamma^2\omega\,\mathrm{sech}^2(x\gamma+\delta_0). \quad (6.10)$$

We separate the variables in the usual manner by letting

$$\omega(x,t)=X(x)T(t). \quad (6.11a)$$

Then

$$T'=\lambda^2T, \quad (6.11b)$$

$$X''+[2\gamma^2\,\mathrm{sech}^2(x\gamma+\delta_0)-\lambda^2]X=0. \quad (6.11c)$$

Obviously

$$T=\exp(t\lambda^2).$$

To find the solution of (6.11c) we note that the x-dependent factor of (3.9) is $\exp\pm\lambda x$. This and (6.2) suggest the solution of (6.11c) to be

$$X(x)=e^{\pm\lambda x}[\pm\lambda-\gamma\tanh(x\gamma+\delta_0)]. \quad (6.12a)$$

This can be verified by direct substitution into (6.11c).

The general solution of (6.11c) is any linear combination of the two solutions given by (6.12a) or

$$X(x)=\lambda[A\cosh\lambda x+B\sinh\lambda x]-\gamma[A\sinh\lambda x+B\cosh\lambda x]\tanh(x\gamma+\delta_0).$$

The constants A and B are to be determined from the boundary conditions (6.1). Since

$$X(0)=\lambda A-\gamma B\tanh\delta_0 \quad \text{and} \quad \gamma\tanh\delta_0=y_F-\alpha_2=u_0,$$

[13] A similar reduction has been made in a heat conduction problem, cf. A. N. Lowan, Phil. Mag. **17**, 849 (1934).

(Eq. 2.11), we have $A = Bu_0/\lambda$ and

$$X(x) = B'\{\lambda[u_0 \cosh\lambda x + \lambda \sinh\lambda x] - \gamma[u_0 \sinh\lambda x + \lambda \cosh\lambda x]\tanh(x\gamma + \delta_0)\}, \quad (6.12b)$$

where B' is a new constant. In order for $X(x)$ to satisfy the second boundary condition, λ must be a root of the transcendental equation (here we have used (2.11) and (2.8))

$$\tanh\lambda l[\lambda^2(u_0+\kappa) - u_0\alpha_1\alpha_2] + \lambda[\lambda^2 + \kappa u_0 - \alpha_1\alpha_2] = 0. \quad (6.13)$$

The general expression for $\omega(x, t)$ is then

$$\omega(x, t) = \sum B_j e^{t\lambda_j^2}[\lambda_j(u_0 \cosh x\lambda_j + \lambda_j \sinh\lambda_j x) - \gamma(u_0 \sinh x\lambda_j + \lambda_j \cosh x\lambda_j)\tanh(x\gamma + \delta_0)], \quad (6.14)$$

where the constants $\{B_j\}$ are to be determined from the initial concentration distribution of A.

We still have the problem of finding $\varphi(x, t)$ when $\omega(x, t)$ is known. This is done by considering (6.2) as a linear differential equation for $\varphi(x, t)$ with $\omega(x, t)$ as a nonhomogeneous term. The solution of this problem and the formula for the B_j's are derived in Appendix II. It is found that

$$\varphi(x, t) = B_0 \cosh(x\gamma + \delta_0)e^{t\gamma^2} + \sum_{j>0} B_j e^{t\lambda_j^2} v_j(x), \quad (6.15)$$

where the λ_j's are solutions of the characteristic Eq. (6.13), the B_j's are given by

$$B_j = \int_0^l \omega(x, 0)X_j(x)dx \Big/ \int_0^l X_j^2(x)dx, \quad (6.16a)$$

with

$$\int_0^l \omega(x, 0)X_j(x)dx = (\gamma^2 - \lambda_j^2) \times \left\{ v_j(l)\varphi_0(l)/\kappa + \int_0^l v_j(x)\varphi_0(x)dx \right\}, \quad (6.16b)$$

$$v_j(x) = u_0 \sinh x\lambda_j + \lambda_j \cosh x\lambda_j, \quad (6.16c)$$

$$\varphi_0(x) = \exp\int_0^x [c_0(x/\psi) - \alpha_2]dx, \quad (6.16d)$$

and

$$\int_0^l X_j^2(x)dx = \tfrac{1}{2}(\lambda_j^2 - \gamma^2) \times \left\{ u_0 + l(u_0^2 - \lambda_j^2) - \kappa\frac{(\lambda_j^2 + \alpha_1\alpha_2)(u_0^2 - \lambda_j^2)}{(\alpha_1^2 - \lambda_j^2)(\alpha_2^2 - \lambda_j^2)} \right\}. \quad (6.16e)$$

The final constant B_0 is chosen to make $\varphi(x, t)$ go continuously into $\varphi_0(x)$ as $t \to 0$. Only one value of x needs to be used for this purpose. For completeness it might be pointed out that

$$v_j(l)/\kappa = \lambda_j(\lambda_j^2 - u_0^2) \times \cosh l\lambda_j / [\lambda_j^2(u_0+\kappa) - u_0\alpha_1\alpha_2]. \quad (6.17)$$

Hence, there is no problem of (6.16b) becoming infinite as $\kappa \to 0$ (i.e., in the case of no production).

When no material is withdrawn from the top of the cascade the expression for $\varphi(x, t)$ becomes considerably simpler. Under this condition $\kappa = 0$, $\alpha_1 = \alpha_2 = \frac{1}{2}$, $u_0 = y_F - \frac{1}{2} = -\sigma$, $\gamma = \frac{1}{2}$, and $\tanh\delta_0 = -2\sigma$. Then the characteristic Eq. (6.13) becomes

$$\sigma \tanh\lambda l = \lambda. \quad (6.18a)$$

An equivalent form is

$$\cosh^2\lambda l = \sigma^2/(\sigma^2 - \lambda^2). \quad (6.18b)$$

These equations and (6.6c) imply that

$$v_j(x) \cosh l\lambda_j = \sigma \sinh\lambda_j(l-x). \quad (6.19)$$

Then

$$\varphi(x, t) = B_0 e^{\frac{1}{4}t} \cosh(\tfrac{1}{2}x + \delta_0) + \sigma \sum (B_j/\cosh l\lambda_j) \exp(t\lambda_j^2) \sinh\lambda_j(l-x). \quad (6.20)$$

We find B_0 by noting that when $x = l$ and $t = 0$

$$\varphi(l, 0) = \varphi_0(l) = \exp\int_0^l [c_0(x/\psi) - \tfrac{1}{2}]dx = B_0 \cosh(\tfrac{1}{2}l + \delta_0).$$

But

$$[\cosh(\tfrac{1}{2}x + \delta_0)/\cosh(\tfrac{1}{2}l + \delta_0)] = [y_F e^{\frac{1}{2}x} + (1-y_F)e^{-\frac{1}{2}x}]/[y_F e^{\frac{1}{2}l} + (1-y_F)e^{-\frac{1}{2}l}].$$

Hence, after the proper reduction of (6.16) we finally obtain

$$\varphi(x, t) = \left\{ \frac{y_F e^{\frac{1}{2}x} + (1-y_F)e^{-\frac{1}{2}x}}{y_F e^{\frac{1}{2}l} + (1-y_F)e^{-\frac{1}{2}l}} \right\} \times \exp\left\{ \tfrac{1}{4}t + \int_0^l [c_0(x/\psi) - \tfrac{1}{2}]dx \right\} + 2\sum_j \frac{\sigma^2 - \lambda_j^2}{[\sigma - l(\sigma^2 - \lambda_j^2)]} \left\{ \frac{\lambda_j\varphi_0(l)}{\lambda_j^2 - \frac{1}{4}} + \int_0^l \varphi_0(y) \sinh\lambda_j(l-y)dy \right\} \times e^{t\lambda_j^2} \sinh(l-x)\lambda_j. \quad (6.21)$$

When $c_0(n) = y_F$ and $P/L\psi = \kappa = 0$ we use

$$\int_0^y (y_F - \tfrac{1}{2})dy = -\sigma y,$$

$$\int_0^l e^{-\sigma y} \sinh\lambda_j(l-y)dy = \lambda_j e^{-l\sigma}/(\sigma^2 - \lambda_j^2),$$

to obtain

$$\varphi(x,t)=\frac{e^{\frac{1}{2}t-\sigma l}[y_F e^{\frac{1}{2}x}+(1-y_F)e^{-\frac{1}{2}x}]}{[y_F e^{\frac{1}{2}l}+(1-y_F)e^{-\frac{1}{2}l}]}$$
$$-8\sum_j \frac{(1-4\sigma^2)\lambda_j \exp(-l\sigma+t\lambda_j^2)\sinh\lambda_j(l-y)}{(4\lambda_j^2-1)\sigma[1-l\sigma(1-\lambda_j^2\sigma^{-2})]}. \quad (6.22)$$

This is exactly the result derived by Majumdar[9] for the corresponding problem in separation by thermal diffusion.

We shall postpone discussion of the application of various solutions of the separation equations here obtained until the next paper of this series. There we shall plot concentration curves as a function of stage number and time for various types of unsteady plant operations. Approximate equations will also be derived to describe these operations. Finally, the exact and simplified approximate solutions of the separation equations will be compared.

APPENDIX I. SOME REMARKS ON LINEARIZATION OF A CLASS OF NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS

Let us consider the general second-order linear partial differential equation in n variables

$$\sum_{ij} A_{ij}f_{x_ix_j}+\sum_i A_i f_{x_i}+Af=0 \quad i,j=1,2,\cdots,n, \quad (I\text{-}1)$$

with $f_{x_j}=\partial f/\partial x_j$, etc. Here the A's may be functions of the independent variables $\{x_j\}$. We introduce a new dependent variable u through

$$f=g(u), \quad (I\text{-}2)$$

where $g(u)$ is some known function of u. Then

$$f_{x_j}=g'(u)u_{x_j},$$
$$f_{x_jx_i}=g''(u)u_{x_j}u_{x_i}+g'(u)u_{x_ix_j},$$

and

$$g''(u)\sum_{ij} A_{ij}u_{x_i}u_{x_j}+g'(u)\sum_{ij}(A_{ij}u_{x_ix_j}+A_j u_{x_i}\delta_{ij})$$
$$+g(u)Au=0. \quad (I\text{-}3)$$

This is in general a nonlinear partial differential equation whose detailed character depends on the choice of the function $g(u)$.

The class of functions $g(u)$ for which a unique inverse $u=g^{-1}(f)$ exists generates a class of nonlinear equations which can be linearized through a transformation in the dependent variable. Of course it hardly needs to be said that the effective use of this linearization process is limited to those nonlinear boundary value problems in which the boundary conditions are also linear after the transformation.

If $u(x_1, x_2, \cdots, x_n)$ is given on a surface

$$\omega(x_1, x_2, \cdots, x_n)=0,$$

then $f=g^{-1}(u)$ can be specified on the same surface. The linearization is useful in this case. Although the boundary conditions required in this paper involved derivatives of u we were fortunate in that they were linearized by the same transformation which linearized the differential equation of interest.

The special form of g, $g(u)=\exp\alpha u$ linearizes all the problems mentioned in Section 3 of this paper. Then

$$\alpha^2\sum_{ij} A_{ij}u_{x_ix_j}+\alpha\sum_{ij}(A_{ij}u_{x_ix_j}+A_j u_{x_j}\delta_{ij})+Au=0. \quad (I\text{-}4)$$

This reduces to the equation of Thomas[7] if $n=2$, $\alpha=c$, $A_{11}=A_{22}=0$, $A_{12}=\frac{1}{2}$, $A_1=A$ and $A_2=B$. It reduces to an equation of the form considered by Majumdar[9] in the general theory of relativity when $n=3$, $A=0$, $\alpha=-\frac{1}{2}$, $A_{ij}=\delta_{ij}$ and $A_j=0$.

The Burgers equation

$$u_t=\mu u_{xx}-u_x u,$$

and our cascade (and Majumdar's thermal diffusion) equation

$$u_t=u_{xx}+2u_x u,$$

are equivalent, as can be verified by letting $t=4\mu\tau$ and $x=-2\mu\xi$ in the Burgers equation (here τ and ξ are the new independent variables). Although (I-5) is not of the form (I-4) as it stands, it is easily transformed to (I-4) by letting $\varphi=\int_0^x u dx$. Then $\varphi_x=u$ and $(\varphi_t)_x=(\varphi_{xx}+\varphi_x^2)_x$. After integrating with respect to x, we obtain $\varphi_t=\varphi_{xx}+\varphi_x^2$ which is the form of (I-4) if one lets $n=2$, $A_{11}=1$, $A_{12}=A_{22}=0$, $\alpha=1$, $A_1=0$, $A_2=-1$ and $A=0$.

It is easy to see that if $g(u)=u^{\alpha+1}$ and $A=0$, then (I-3) is a "quadratic" equation

$$\alpha\sum_{ij} A_{ij}u_{x_i}u_{x_j}+\sum_{ij}(A_{ij}uu_{x_ix_j}+A_i uu_{x_j}\delta_{ij})=0.$$

The reader can easily construct many other equations from (I-3).

Other classes of nonlinear equations analogous to (I-3) can be generated by transformations such as $f=hg(ku)$ where h and k are known functions of $(x_1, x_2, \cdots, x_n)$ and g is a known function of ku. One can proceed in a similar manner with simultaneous linear equations in several dependent variables.

The authors are indebted to Professor C. Arf for a useful discussion concerning this appendix.

APPENDIX II. DERIVATION OF EQUATION (6.15)

In Section 6 we gave the final expression for $\varphi(x,t)$ without proof. Here we shall start with (6.14) and supply the required steps in the derivation of (6.15).

The function $\varphi(x,t)$ can be obtained from (6.2) in terms of $\omega(x,t)$. Then

$$\varphi(x,t)=\left[a(t)+\int_0^x d\xi\,\omega(\xi,t)\exp\left(-\int_0^\xi f(y)dy\right)\right]$$
$$\times\exp\int_0^x f(y)dy, \quad (II\text{-}1)$$

where $a(t)$ is an arbitrary function of t (the integration "constant"). We choose $a(t)$ so that $\varphi(x, t)$ satisfies $\varphi_{xx}=\varphi_t$. From (II-1)

$$\varphi_t-\varphi_{xx}=\left[\exp\int_0^x f(y)dy\right]\{a_t-a(f_x+f^2)-\omega_x(0, t)\}.$$

Hence to make $\varphi_t=\varphi_{xx}$, we require

$$a_t=a(f_x+f^2)+\omega_x(0, t). \tag{II-2}$$

Since $f(x)=\gamma\tanh(\gamma x+\delta_0)$, $f_x+f^2=\gamma^2$ we have

$$a(t)=a_0\exp(t\gamma^2)+\exp(t\gamma^2)\int_0^t \omega_x(0, \tau)\exp(-\tau\gamma^2)d\tau,$$

where a_0 is in a constant. Then

$$\varphi(x, t)=\operatorname{sech}\delta_0\cosh(x\gamma+\delta_0)\Big\{a_0\exp(t\gamma^2)$$

$$+\exp(t\gamma^2)\int_0^t \omega_x(0, \tau)\exp(-\tau\gamma^2)d\tau$$

$$+\int_0^x \omega(\xi, t)\cosh\delta_0\operatorname{sech}(\xi\gamma+\delta_0)d\xi\Big\}. \tag{II-3}$$

To eliminate $\omega_x(0, \tau)$ and $\omega(\xi, t)$ we use (6.14). Then from the definition of u_0, δ_0, $v_j(x)$ (Eq. 6.16c), etc.

$$\omega_x(0, t)=\sum\lambda_jB_j\exp(t\lambda_j^2)(\lambda_j^2-\gamma^2),$$

$$\int_0^t \omega_x(0, \tau)\exp(-\tau\gamma^2)d\tau=\sum\lambda_jB_j\{\exp[(\lambda_j^2-\gamma^2)t]-1\},$$

$$\int_0^x \omega(\xi, t)\operatorname{sech}(\xi\gamma+\delta_0)d\xi$$

$$=\sum B_j\exp(t\lambda_j^2)\{v_j(x)\operatorname{sech}(x\gamma+\delta_0)-\lambda_j\operatorname{sech}\delta_0\}.$$

Hence

$$\varphi(x, t)=\operatorname{sech}\delta_0\cosh(x\gamma+\delta_0)(a_0-\sum\lambda_jB_j)e^{\gamma^2t}$$
$$+\sum B_j\exp(t\lambda_j^2)v_j(x).$$

Now $\sum\lambda_jB_j$ is independent of x. Therefore we lump all constants into B_0 and find

$$\varphi(x, t)=B_0\cosh(x\gamma+\delta_0)e^{t\gamma^2}$$
$$+\sum_j B_j\exp(t\lambda_j^2)v_j(x). \tag{II-4}$$

Unfortunately, since the functions $\{v_j(x)\}$ do not form an orthogonal set we must return to the initial conditions on $\omega(x, t)$ to determine the constants $\{B_j\}$. The characteristic functions

$$X_j(x)=v_j'(x)-f(x)v_j(x), \tag{II-5}$$

of (6.11) do form an orthogonal set if the λ_j's are simple. Hence, from (6.14)

$$B_j=\int_0^l \omega(x, 0)X_j(x)dx\Big/\int_0^l X_j^2(x)dx. \tag{II-6}$$

Now (6.2) implies

$$I=\int_0^l \omega(x, 0)X_j(x)dx$$

$$=\int_0^l[\varphi_xv_j'(x)-f(\varphi v_j)_x+f^2v_j\varphi]dx$$

$$=\varphi v_j'(x)\Big]_0^l-(f\varphi v_j)\Big]_0^l$$

$$-\int_0^l \varphi v_j''(x)dx+\int_0^l \varphi v_j(f_x+f^2)dx.$$

Since $f_x+f^2=\gamma^2$, $v_j''=\lambda_j^2v_j$, $v_j(0)=\lambda_j$, $v_j'(0)=u_0\lambda_0$, $f(0)=0$, we have

$$I_1=[\varphi(l, 0)v_j'(l)-f(l)\varphi(l, 0)v_j(l)]$$
$$+(\gamma^2-\lambda_j^2)\int_0^l v_j\varphi(x, 0)dx.$$

The transcendental equation (6.13) implies that

$$\kappa[v_j'(l)-v_j(l)f(l)]=(\gamma^2-\lambda_j^2)v_j(l).$$

Hence,

$$I_1=(\gamma^2-\lambda_j^2)\Big\{v_j(l)\varphi_0(l)/\kappa+\int_0^l v_j(x)\varphi_0(x)dx\Big\}. \tag{II-7}$$

One can derive Eq. (6.16e) for $\int_0^l X_j^2dx$ in a similar manner.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 2 FEBRUARY, 1952

An Iterative Method for the Solution of Eigenvalue Problems

V. J. BERRY* AND C. R. DE PRIMA†
California Institute of Technology, Pasadena, California

(Received July 13, 1951)

A simple iterative procedure is developed for the determination of eigenvalues and eigenfunctions associated with the solution of Sturm-Liouville problems in a finite interval. Both the frequency and the displacement for a given mode may be determined with an accuracy which is independent of the accuracy involved in the calculation of other modes. Convergence of the iteration process is rapid, and the successive approximations to a given eigenvalue are shown to form a monotone sequence. The method is particularly useful when the coefficients of the differential equation are not expressed in analytical form.

INTRODUCTION

CERTAIN boundary value problems of interest in physics and engineering give rise to Sturm-Liouville systems for which even the approximate determination of eigenvalues and eigenfunctions is difficult. Although in such instances the calculation of the lowest eigenvalue and the first eigenfunction is usually simple, the methods for the determination of higher eigenvalues and eigenfunctions are generally tedious. For example, the method of "sweeping"[1] requires the choice of trial functions which are already orthogonal to the lower modes, while the Rayleigh-Ritz procedure involves the solution of many simultaneous equations or the determination of the roots of a polynomial of large degree if very good approximations to the eigenvalues are desired. In fact, most methods usually require the simultaneous solution of a number of algebraic equations.[2]

A simple, iterative scheme of calculation is presented here which, in principle, yields all modes and the corresponding eigenvalues with any desired accuracy, and which does not require the choice of a trial function which is orthogonal to the lower modes. The essential feature of the method is that it requires the direct solution of a related initial value problem, rather than the solution of the actual boundary value problem itself. A special advantage is that the accuracy of determination of a given eigenvalue and the corresponding eigenfunction is independent of the accuracy involved in computation of all other modes, and depends only upon the accuracy with which one obtains a solution of the related initial value problem. The calculations involve the numerical integration of a second-order ordinary differential equation, and are quite routine and simple in their nature; the method being particularly useful when a desk calculator is available, and when it is desired to calculate only the first two or three modes. In the following paragraphs the principle of the iteration method will first be developed, and then a numerical example will be chosen to illustrate the method.

THE ITERATION METHOD‡

Consider

$$L(u)\equiv(pu')'-qu=-\lambda wu \quad 0<x<1$$

$$u(0)=0, \quad \int_0^1 wu^2dx=1 \tag{1}$$

$$u'(1)=0, \tag{2}$$

with $p(x)$, $p'(x)$, $q(x)$, and $w(x)$ continuous in $0\leq x\leq 1$, and $p(x)>0$, $w(x)>0$ in $0<x\leq 1$.§ Then there exists[3] a countable number of eigenvalues $\lambda_1, \lambda_2, \cdots\lambda_n, \cdots$ and corresponding eigenfunctions $u_1(x), u_2(x), \cdots u_n(x), \cdots$ such that u_n has precisely $n-1$ zeros in $0<x\leq 1$.

If $u(x, \lambda)$ is a solution of (1), then upon multiplication of (1) by u and integration over the region $0\leq x\leq 1$, there results

$$\lambda=\int_0^1 (pu'^2+qu^2)dx-p(1)u(1, \lambda)u'(1, \lambda). \tag{3}$$

Lemma

$$D(\lambda)\equiv\int_0^1 (pu'^2+qu^2)dx$$

is a monotone increasing function of λ for $\lambda>\lambda_1$, where $u(x, \lambda)$ satisfies (1).

* Present address, Stanolind Oil and Gas Company, Tulsa, Oklahoma.

† Presently on leave with the ONR, Mathematics Branch, Washington 25, D.C.

[1] L. Beskin and R. M. Rosenberg, J. Aeronaut. Sci. **13**, 597 (1946).

[2] See, for example, L. Collatz, *Eigenwertprobleme* (Chelsea Publishing Company, New York, 1948), p. 264.

‡ Although the following discussion applies to the solution of symmetrical problems, the extension of the method to unsymmetrical cases involving general, homogeneous boundary conditions should be obvious.

§ The prime will be used throughout to denote differentiation with respect to x.

[3] For a complete discussion of this theory see E. L. Ince, *Ordinary Differential Equations* (Dover Publications, New York), first American edition, Chapter X.

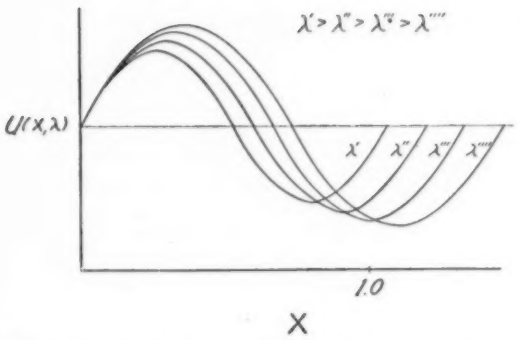


FIG. 1. Variation of the function $u(x, \lambda)$ where u is normalized and satisfies the equation $(pu')'-qu+\lambda wu=0$ with the initial condition $u(0)=0$.

Proof

The derivative $dD/d\lambda$ exists and is given by the following.

$$dD/d\lambda=2\int_0^1[pu'(du'/d\lambda)+qu(du/d\lambda)]dx$$

$$=2\int_0^1(du/d\lambda)[-(pu')'+qu]dx$$

$$+2p(1)u'(1,\lambda)du(1,\lambda)/d\lambda$$

$$dD/d\lambda=2\lambda\int_0^1 wu(du/d\lambda)dx$$

$$+2p(1)u'(1,\lambda)du(1,\lambda)/d\lambda. \quad (4)$$

But by differentiation of the normalization condition stated in Eq. (1),

$$\int_0^1 wu(du/d\lambda)dx=0. \quad (5)$$

Therefore

$$dD(\lambda)/d\lambda=2p(1)u'(1,\lambda)du(1,\lambda)/d\lambda. \quad (6)$$

In Fig. 1 are shown several functions $u(x, \lambda)$ for λ near λ_2. The general behavior of $u(x, \lambda)$ is known.[4] Upon inspection of Fig. 1 the truth of the lemma is intuitively obvious. This follows directly from the fact that $u(1, \lambda)$ has either a relative maximum or minimum, if and only if $u'(1, \lambda)=0$. A simple consideration shows that $u'(1, \lambda)$ and $du(1, \lambda)/d\lambda$ are both of the same sign. Thus, the only stationary values of $D(\lambda)$ are inflection points.

Theorem 1

The eigenvalues of (1) and (2) are stationary values of $D(\lambda)$.

[4] See reference 3, pp. 231–233.

Proof

Since $p(1)>0$ by hypothesis, $dD/d\lambda$ vanishes only at the zeros of $u'(1, \lambda)$ and $du(1, \lambda)/d\lambda$. But since $u'(1, \lambda)$ and $du(1, \lambda)/d\lambda$ are oscillating functions of λ and are both of the same sign by the foregoing lemma, $du(1, \lambda)/d\lambda=0$, if and only if $u'(1, \lambda)=0$. Therefore, $D(\lambda)$ has an inflection point only when $u'(1, \lambda)=0$. But this is precisely the condition that λ be an eigenvalue of (1) and (2), and hence the theorem is proved.

Figure 2 depicts the behavior of $D(\lambda)$ for the special case $p=1$, $q=0$, and $w=1$. Under these conditions (1) becomes

$$u''+\lambda u=0$$

$$u(0)=0, \quad \int_0^1 u^2dx=1. \quad (7)$$

The eigensolutions are, for $u'(1)=0$, $u(x, \lambda_n)\equiv u_n(x)=\sqrt{2}\sin(n-\frac{1}{2})\pi x$ with $\lambda_n=[(n-\frac{1}{2})\pi]^2$.

Now let $\lambda_n^{(1)}$ be a first approximation to λ_n such that $\mu_{n-1}<\lambda_n^{(1)}<\mu_n$ where μ_{n-1} and μ_n are the $n-1$st and nth eigenvalues of (1) with the conditions $u(0)=u(1)=0$.|| There are several methods for estimating a suitable value for $\lambda_n^{(1)}$. If $w(0)>0$, one method is to compare (1) with the eigenvalues of the following problem

$$M(v)\equiv(\bar{p}v')'-\bar{q}v=-\lambda\underline{w}v$$

$$v(0)=0, \quad \int_0^1 \underline{w}v^2dx=1 \quad (8)$$

$$v'(1)=0, \quad (9)$$

where $\bar{p}=\max_{0\le x\le 1} p(x)$, $\bar{q}=\max_{0\le x\le 1} q(x)$, and $\underline{w}=\min_{0\le x\le 1} w(x)$. This method is useful for the estimation of lower eigenvalues, whereas the estimation of the higher eigenvalues may be facilitated by a knowledge of the values of the

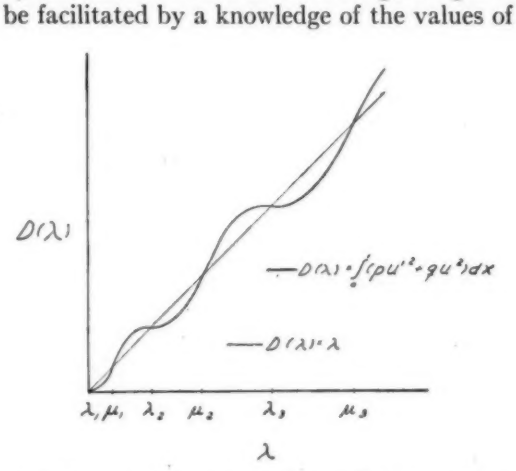


FIG. 2. Behavior of the function $D(\lambda)\equiv\int_0^1(pu'^2+qu^2)dx$ where u is normalized and satisfies the equation $(pu')'-qu+\lambda wu=0$ with the initial condition $u(0)=0$.

|| A separate consideration is necessary for λ_1.

lower ones. Another method of approximation is to use the asymptotic expression for the eigenvalues.[5]

$$\lambda_n=\left[(n-\tfrac{1}{2})\pi\Big/\int_0^1\left(\frac{w}{p}\right)^{\frac{1}{2}}dx\right]^2+O(1). \tag{10}$$

The Rayleigh quotient[6] will also yield an estimate of λ_n even if a trial function is chosen which is not orthogonal to the lower modes. In general, however, the best aid in selecting a value for $\lambda_n^{(1)}$ is a little familiarity with the differential equation itself and the boundary value problem under consideration.

Now, if $\mu_{n-1}<\lambda_n^{(1)}<\mu_n$, then $u(x, \lambda_n^{(1)})$ will have precisely $n-1$ zeros in $0<x\leq 1$ where $u(x, \lambda_n^{(1)})$ satisfies (1). Equation (3) becomes for these conditions

$$\lambda_n^{(1)}=D(\lambda_n^{(1)})-p(1)u(1, \lambda_n^{(1)})u'(1, \lambda_n^{(1)}). \tag{11}$$

Now define

$$\lambda_n^{(2)}\equiv D(\lambda_n^{(1)}). \tag{12}$$

Then

$$\lambda_n^{(2)}=\lambda_n^{(1)}+G(\lambda_n^{(1)}), \tag{13}$$

where

$$G(\lambda)\equiv p(1)u(1, \lambda)u'(1, \lambda). \tag{14}$$

Or continuing the iteration process

$$\lambda_n^{(K+1)}=\lambda_n^{(K)}+G(\lambda_n^{(K)}). \tag{15}$$

Theorem 2

The sequence of approximations $\lambda_n^{(K)}$ converges monotonically to λ_n where $\lambda_n^{(K)}\equiv D(\lambda_n^{(K-1)})$.

Proof

By consideration of Fig. 2 the proof is obvious. If $\mu_{n-1}<\lambda_n^{(1)}<\lambda_n$ then $\lambda_n^{(K+1)}\equiv D(\lambda_n^{(K)})>\lambda_n^{(1)}$. But $\lambda_n^{(K+1)}\equiv D(\lambda_n^{(K)})<\lambda_n$ because of the monotone character of $D(\lambda)$. Similarly for $\lambda_n<\lambda_n^{(1)}<\mu_n$. Now, since the $\lambda_n^{(K)}$ from a monotone sequence they have a limit point λ_n'. Thus $\lambda_n'\leftarrow\lambda_n^{(K)}\equiv D(\lambda_n^{(K-1)})\rightarrow D(\lambda_n')$. Therefore $\lambda_n'=\lambda_n$. By the continuity of u, $u(x, \lambda_n^{(K)})\rightarrow u(x, \lambda_n)\equiv u_n(x)$.

NUMERICAL EXAMPLE

In the following example we have chosen a boundary value problem whose exact solution is known so that the iteration procedure might be compared step by step with the analytical solution. The present method is by no means restricted, however, even to those cases in which $p(x)$, $q(x)$, and $w(x)$ are expressed in analytical form, but is readily adaptable to the solution of problems in which these functions are given in the form of tables.

[5] See reference 3, pp. 270–273.

[6] See reference 2, p. 70.

TABLE I Successive approximations to $u_4(x)$ and λ_4 where $u_4(x)$ is normalized and satisfies the equation $(xu')-(1/x)u+\lambda xu=0$ with the conditions $u_4(0)=0$ and $u_4'(1)=0$. λ_4 is the fourth root of the equation $J_1'[(\lambda)^{\frac{1}{2}}]=0$.

K	1	2	3	4	∞
$\lambda_4^{(K)}$	120.905	131.870	136.869	137.031	137.031
x	$u_4^{(1)}(x)$	$u_4^{(2)}(x)$	$u_4^{(3)}(x)$	$u_4^{(4)}(x)$	$u_4(x)$
0	0	0	0	0	0
0.1	2.8378	2.9471	2.9838	2.9855	2.9855
0.2	3.3513	3.2872	3.2391	3.2381	3.2381
0.3	1.3322	0.9758	0.8102	0.8051	0.8051
0.4	−1.2184	−1.5503	−1.6720	−1.6746	−1.6746
0.5	−2.0583	−1.9404	−1.8460	−1.8431	−1.8431
0.6	−0.7567	−0.2307	0.0062	0.0153	0.0153
0.7	1.0888	1.4590	1.5634	1.5663	1.5663
0.8	1.5926	1.3360	1.1531	1.1462	1.1462
0.9	0.4180	−0.2424	−0.5199	−0.5281	−0.5281
1.0	−1.0616	−1.3840	−1.4194	−1.4194	−1.4194

In Eq. (1) let $p(x)=x$, $q(x)=1/x$, and $w(x)=x$. Then (1) becomes

$$(xu')'-\frac{1}{x}u+\lambda xu=0 \quad 0<x<1$$

$$u(0)=0, \quad \int_0^1 xu^2dx=1 \tag{16}$$

$$u'(1)=0. \tag{17}$$

The exact solution of (16) and (17) is

$$u_n(x)=J_1[(\lambda_n)^{\frac{1}{2}}x]/N_n, \tag{18}$$

where λ_n is determined from the expression

$$J_1'[(\lambda_n)^{\frac{1}{2}}]=0. \tag{19}$$

The normalizing constant has the value

$$N_n^2=(\lambda_n-1)J_1^2[(\lambda_n)^{\frac{1}{2}}]/2\lambda_n. \tag{20}$$

Now suppose it is desired to determine the 4th eigenvalue and corresponding eigenfunction of (16) and (17) according to the present iteration method. We calculate $\lambda_4^{(1)}$ from Eq. (10) with $n=4$.

$$\lambda_4^{(1)}=\left[7\pi/2\int_0^1\left(\frac{x}{x}\right)^{\frac{1}{2}}dx\right]^2=120.905. \tag{21}$$

Substitution of this value of $\lambda_4^{(1)}$ into (16) followed by integration¶ yields the results shown in Table I. One may immediately determine from these results $G(\lambda_4^{(1)})=10.965$. Hence

$$\lambda_4^{(2)}=\lambda_4^{(1)}+10.965=131.870. \tag{22}$$

The next integration yields $G(\lambda_4^{(2)})=4.999$, resulting in $\lambda_4^{(3)}=136.869$ which is within 0.1 percent of the true

¶ In practice, any of the standard methods of integration of linear, second-order differential equations are applicable. Formulas for numerical solution of second-order differential equations are given, for example, by W. E. Milne, *Numerical Calculus* (Princeton University Press, Princeton, New Jersey, 1949), Chapter V, and by others.

eigenvalue $\lambda_4 = 137.031$. Table I summarizes the results of the iterative process carried out to $K=4$. The corresponding $u(x)$ is compared with the exact solution of (16) and (17). The approximation $\lambda_4^{(4)}$ agrees exactly with the eigenvalue $\lambda_4 = 137.031$ to the number of decimals given.

CONCLUSION

An iterative method has been developed for the solution of eigenvalue problems associated with Sturm-Liouville systems. The convergence of the sequence $\lambda_n^{(K)}$ is shown to be monotone, and is, in fact, reasonably rapid, as is demonstrated in the numerical example.

The eigenvalues and eigenfunctions may be determined subject to the accuracy with which Eq. (1) may be integrated for a given value of λ. In the cases in which $p(x)$, $q(x)$, and $w(x)$ are expressed in analytical form, Eq. (1) may be integrated exactly as an infinite series in x and λ. In such situations the quantity $G(\lambda)$ may be expressed in analytic form. The present method is even more appropriate, however, in those cases in which $p(x)$, $q(x)$, and $w(x)$ exist only in the form of a table of values given as a function of the argument x, a situation which is common to many problems of physical interest. Under these conditions (1) may be solved by numerical integration and the function $G(\lambda) \equiv p(1)u(1, \lambda)u'(1, \lambda)$ determined numerically for each successive approximation.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 2 FEBRUARY, 1952

Phenomena Associated with the Flight of Ultra-Speed Pellets.* Part II. Spectral Character of Luminosity†

W. C. White, John S. Rinehart, and William A. Allen
Michelson Laboratory, United States Naval Ordnance Test Station, Inyokern, China Lake, California

(Received July 16, 1951)

The spectroscopic character of the luminous trails generated by ultra-speed pellets has been investigated. A total of seven spectrograms have been measured and reduced. Three of these were of 5.0 km/sec aluminum pellets; two were of 5.9 km/sec magnesium pellets; one was of a 6.0 km/sec magnesium-lithium-aluminum pellet; and one was of a titanium pellet whose velocity was estimated at about 4.5 km/sec.

The pronounced features of the spectrograms are: (1) AlO bands in the aluminum spectra, (2) Mg lines and MgO bands at the magnesium spectra, (3) strong Li lines and a pronounced continuum in the magnesium-lithium-aluminum spectra, and (4) TiO bands in the titanium spectra. Other features of the spectra are: (1) prominent AlI lines at λλ3944 and 3961 and a relatively weak continuum in the aluminum spectra, (2) a stronger continuum in the magnesium spectra, and (3) relatively weak MgO bands in the Mg-Li spectra.

The time resolved spectra show that the AlO persists in the excited state twice as long as the aluminum. The aluminum oxidizes almost completely leaving AlO in an excited state which continues to emit light after the aluminum is burned.

INTRODUCTION

In general, an ultra-speed body moving through air is followed by an extended luminous trail. Previous spectroscopic observation of such trails has been confined largely to the studies of meteors. Excellent discussions of the meteor observations have been presented by Millman[1] and Russell.[2] Meteor spectra consist of the lines of ionized metals.

R. W. Wood[3] has reported spectrograms of the metallic jets formed by cavity charges. In this case, the identified lines could be associated with the material of the metal liners. Numerous other unpublished data obtained both at the U. S. Naval Ordnance Test Station and at other laboratories are in substantial agreement with Wood's observations.

Previous studies have been made under difficulties. Observations on meteors cannot be made under controlled laboratory conditions. Investigations using conventional shaped charges suffer from two disadvantages: (1) it is always necessary to observe a group of fragments and (2) the initial shapes of the fragments are not known.

In the present study it has been possible to fire individual ultra-speed metallic pellets under more or less controlled laboratory conditions. Excellent spectrograms of the luminous trails have been obtained. It is possible to infer certain facts regarding the origin of the luminosity of the pellet's trail. This part of the report describes in detail the observed spectra and discusses the inferences that have been drawn.

EXPERIMENTAL ARRANGEMENTS

Several arrangements have been used to obtain spectrograms. R. W. Wood in his experiments with cavity charges used the luminous metallic jet as a line source. Many excellent spectrograms have been obtained in a similar manner. An example of such a spectrogram is reproduced in Fig. 1. This method has the advantage

* Work supported by Armament Branch, ONR.
† Part I of this series appeared in January 1952, J. Appl. Phys.
[1] P. M. Millman, Harvard Coll. Obs. Ann. **82**, 131 (1932).
[2] J. A. Russell, Contrib. Meteoritical Soc. **4**, 250 (1950).
[3] R. W. Wood, Phys. Rev. **71**, 471(A) (1947).

of simplicity in the arrangement of the experiment since the luminous trail is used in lieu of a slit. Many other spectrograms were taken with a crudely constructed field spectrograph. The spectrograph consisted of a 5-ft box with an adjustable slit at one end and a camera at the other end. This spectrograph was placed close to the trajectory of the pellet. The way in which the box was used is evident in Fig. 1 of Part I.

All of the spectra, except the time-resolved spectra, which will be discussed later, have been obtained with a Combat Graphic Camera used in conjunction with a transmission grating. The camera lens was an f:4.7 Ektar, 127 mm focal length. The grating was a Wood's replica with 14,438 lines per inch. A collimating achromatic lens of focal length 55 in. was placed in front of the grating when the box spectrograph was used. A General Radio High-Speed, continuously-moving film camera with a f:2.9, 100 mm focal length lens was used to obtain the time-resolved spectra. The camera, fitted with a 60-degree flint prism, was placed at the open end of the box. The film was moved at a velocity of about 35 ft/sec, parallel to the slit. The slit was about 4 in. long. For optimum results, the slit width was set at about 0.8 mm. A mercury vapor lamp was used to furnish comparison spectra.

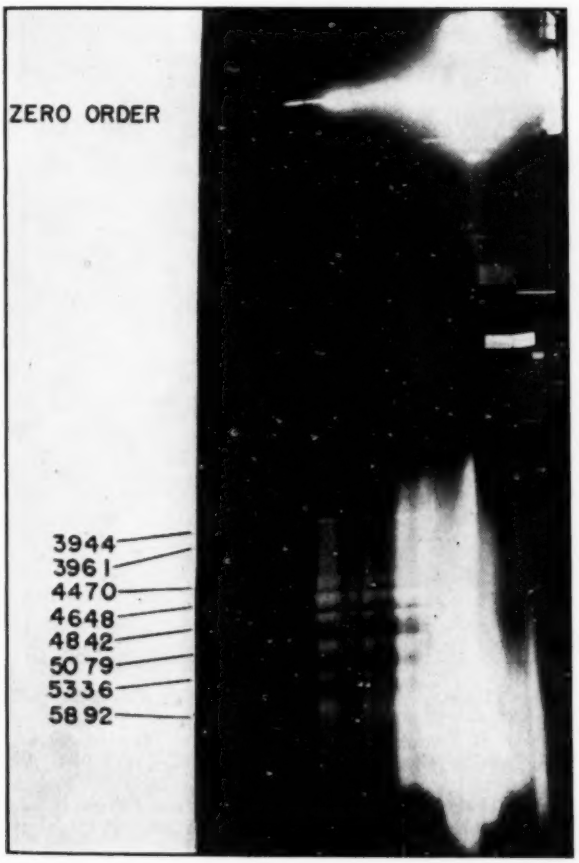


FIG. 1. Spectrogram of the luminous trail generated by a 5.0 km/sec aluminum pellet.

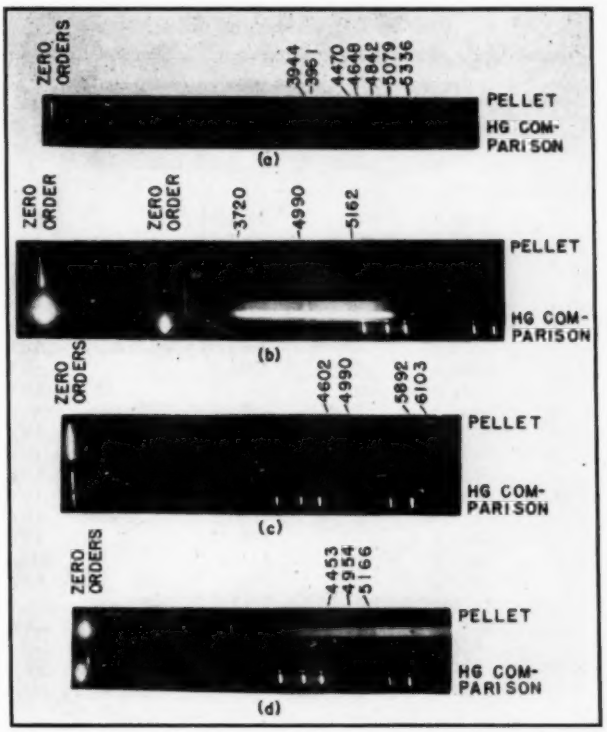


FIG. 2. Spectrograms obtained with the box spectrograph. (a) Aluminum—5.0 km/sec; (b) Magnesium—5.9 km/sec; (c) Magnesium-lithium—6.0 km/sec; (d) Titanium—approximately 4.5 km/sec.

MEASUREMENT AND REDUCTION

A total of seven spectrograms have been measured and reduced. Three of these were of aluminum pellets, two were of magnesium pellets, one was of a magnesium-lithium-aluminum pellet,[4] and one was of a titanium pellet. A typical spectrogram of each pellet is reproduced in Fig. 2.

Each spectrogram was measured in a conventional manner. Only first-order spectra were used. All lines were measured from the zero order as a reference point for both the pellet spectra and the comparison spectra. The measures were reduced in the usual way by application of a quadratic equation.[5] As is evident from Fig. 2, the zero order of the comparison spectra and of the pellet spectra are not colinear on all the plates. The displacement was caused by slightly different orientations of the camera during the two exposures. A correction was applied for this displacement.

SPECIAL CHARACTERISTICS

The data have been grouped in Table I with respect to pellet material and wavelength. The first column gives the number of plates on which the line was observed. The measured value of each wavelength is listed in the second column. When the line was observed on

[4] The composition of this pellet is given in Part I.

[5] R. A. Sawyer, *Experimental Spectroscopy* (Prentice-Hall, Inc., New York, 1944), p. 238.

TABLE I. Wavelengths of lines and bands in the spectra of the trails of ultra-speed pellets.

Number of plates on which line was measured	Observed wavelength A	Visual estimate of intensity[a]	Line identification Accepted wavelength A		Source
Pellet material: 2S—O *Aluminum alloy*					
3	3943.9	6	3944.54		Al
3	3961.5	7	3961.54		Al
2	4471.2	2	4470.5		AlO
3	4490.9	2	4494.0	band[b]	AlO
3	4514.8	1	4516.3		AlO
3	4647.1	7	4648.2		AlO
3	4671.5	4	4672.0	band	AlO
3	4693.3	2	4694.6		AlO
3	4841.2	10	4842.1		AlO
3	4866.2	3	4866.1	band	AlO
3	5076.5	7	5079.3		AlO
3	5098.7	5	5102.1		AlO
3	5120.8	3	5123.3	band	AlO
1	5139.8	1	5142.9		AlO
1	5155.2	1	5160.8		AlO
1	5330.2	2	5336.9		AlO
1	5342.8	3	5341.07		Mn
2	5352.3	1	5358.1	band	AlO
3	5371.0	2	5377.4		AlO
2	5391.8	3	5394.8		AlO
3	5452.9	7	5460.73		Hg
2	5892.5	5	5892.62		Na
1	6000.8	1	6017.31		Mn
Pellet material: *Magnesium*					
1	3256.6	2	several blended lines		Fe
1	3279.3	5	several blended lines		Fe
1	3720.3	8	several blended lines		Fe
1	3836.7	8	3833.55		Mg
1	4364.9	3	4351.9		Mg
2	4820.1 to 4893.6	4	4818.5 to 4949.5	band	MgO
2	4983.7 to 4992.7	9	4985.9 to 5007.3	band	MgO
			5167.33		
1	5162.4	4	5172.68	blend	Mg
			5183.60		
2	5880.7	4	5892.6		Na
Pellet material: *Magnesium-Lithium-Aluminum alloy*					
1	3906.4	2	3915		Li
1	4604.2	4	4602.29		Li
			4962.1		MgO
	4961.6		4971.9		Li
1	to	3	4974.5	band	MgO
	5003.2		4985.9		MgO
			4996.7		MgO
			5007.3		MgO
1	5887.1	4	5892.95		Na
1	6099.6	10	6103.6		Li
Pellet material: *Titanium*					
1	4453.7	4	4462.1	band	TiO
1	4936.8	6	4954.6	band	TiO
1	5154.5	8	5166.9	band	TiO

[a] Intensities of lines are indicated by the numbers 1 to 10, the latter the higher intensity.

[b] The rotational structure associated with the individual bands could not be resolved.

more than one plate, the average value is given. A visual estimate of the relative intensity of each line is listed in the third column. The line structures were identified as molecular bands[6] and atomic lines.[7–9] The identified line and the source are listed in the last column of the table.

Three types of emission spectra, line, band, and continuum, are present in all the spectrograms. Much of the titanium spectrum was obscured by radiation from the products of explosion.

The most pronounced features of the spectrograms are: (1) AlO bands in the aluminum spectra, (2) Mg lines and MgO bands in the magnesium spectra, (3) strong Li lines and a pronounced continuum in the magnesium-lithium-aluminum spectra, and (4) TiO bands in the titanium spectra. Other features of the spectra are: (1) prominent AlI lines at λλ3944 and 3961 and a relatively weak continuum in the aluminum spectra, (2) relative to the lines and bands the intensity of the continuum in the magnesium spectra is stronger than in the case of aluminum, and (3) relatively weak MgO bands in the Mg-Li spectra.

The yellow sodium doublet appears in the magnesium and aluminum spectra. It has approximately the same intensity in both spectra. The source of the sodium may be the soil of this area which has a high alkali content.

The lines at λλ5452.9 and λλ6000.8 correspond with

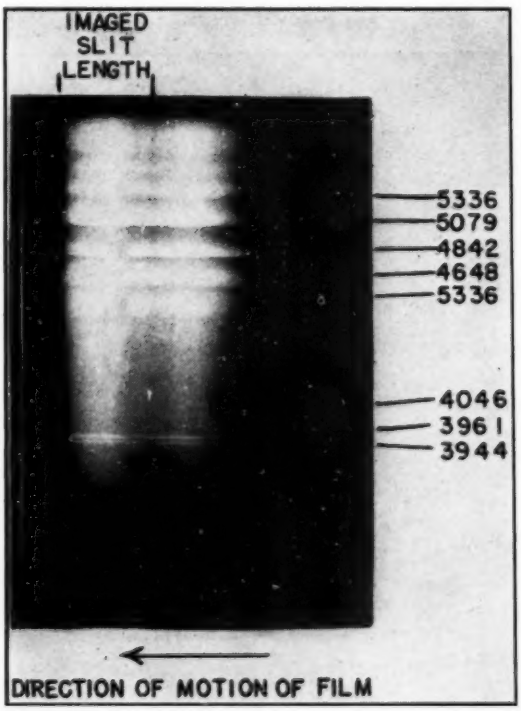


FIG. 3. Time resolved spectra of 5.0 km/sec aluminum pellet.

[6] R. W. B. Pearse and A. G. Gaydon, *The Identification of Molecular Spectra* (Chapman and Hall, Ltd., London, 1950), second edition.

[7] C. D. Hodgman, *Handbook of Chemistry and Physics* (Chemical Rubber Publishing Company, Cleveland, Ohio, 1948), thirtieth edition.

[8] Massachusetts Institute of Technology, *M.I.T. Wavelength Tables* (John Wiley and Sons, Inc., New York, 1939).

[9] C. E. Moore, Contrib. from Princeton Univ. Obser. **20**, (1945).

TABLE II. Persistence of the burning trail of the aluminum pellets.

Wavelength A	Source	Persistence on individual plates in μ sec
3944	Al	74
		74
		74
		145[a]
4842	AlO	145
		145
		145
		221[a]

[a] These values of persistence are anomalous, but the difference in the times of persistence of the band and line is very nearly the same as that for the other pellets.

those of mercury, λλ5460.73, and manganese, λλ6017.31, respectively. These are very strong lines in the spectra of the two elements and would be expected to appear even if the elements were present in minute quantities. In addition, the 2S−O aluminum alloy used was known to have 1.2 percent manganese as a basic constituent. A faint line appeared near the aluminum doublet on a number of the spectrograms. This line, which was too weak to measure on a comparator, is not listed in Table I. However, a rough measurement made with a millimeter rule established the wavelength as λλ4049, a value close to a strong mercury line at λλ4046. A few iron lines were found in the blue region of the magnesium spectra.

TIME-RESOLVED SPECTRA

Several time-resolved spectra, a typical example of which is reproduced in Fig. 3, were obtained with the General Radio High-Speed Camera. The film was moved parallel to the spectrograph slit at a velocity of approximately 35 ft/sec. The persistent times of the AlI lines and the AlO bands were compared by measuring the lengths of the lines and of the strongest AlO band, λλ4842. A correction was made for the length of the image of the slit. Table II gives the wavelength and persistence of each line on the individual plates.

The lines and the bands seem to appear simultaneously on the film. The time resolution, however, of the General Radio Camera is not high since the writing speed was only 0.01 mm per microsec. The luminous trail persists 0.1 to 0.2 millisecond. This persistence is substantially in agreement with that observed on streak photographs. The aluminum oxide persists about twice as long as the aluminum.

DISCUSSION

The spectra of the trail of ultra-speed pellets depends on the composition of the pellet. There is no evidence for the excitation of the O_2 and N_2 bands in the visual region of the spectra of the pellets. The dependence of the characteristics of the spectra on velocity has not been explored.

The trail of an ultra-speed pellet is luminous because of burning and excitation of material that is ablated from the surface of the pellet. The impact of air molecules on a pellet raises its surface temperature to a point where the metal either melts or vaporizes. This ablated material then collides with air molecules. Oxidation results.

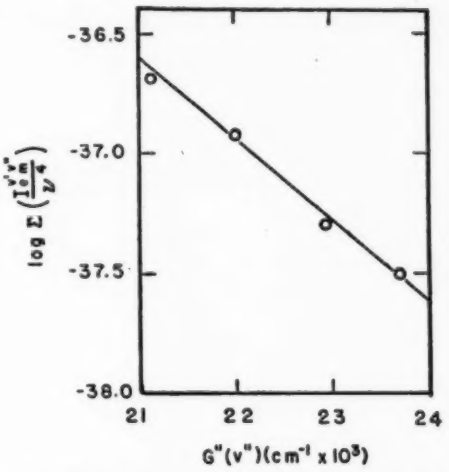


FIG. 4. Plot of data used to obtain the effective vibrational temperature of the trail of an aluminum pellet.

In the case of aluminum, a temperature of the luminous trail has been calculated from visual estimates of the relative intensities of the bands. This temperature has been calculated by the relation[10] where v is the

$$\log \sum_{v''} \frac{I_{em.}{}^{v'v''}}{\nu_{v'v''}{}^{4}} = C_1 - \frac{hc}{kT} G''(V''),$$

quantum number of the vibrational state, $I_{em.}{}^{v'v''}$ is the intensity of the band emitted by the transition $v''\rightarrow v'$, C_1 is a constant, $G''(V'')$ is the energy of the vibrational slate, h is Planck's constant, c is the velocity of light, k is Boltzmann's constant, and T is the temperature in degrees Kelvin. The summation is taken for all observed bands that originate from the same energy level. A plot of the present data, Fig. 4, yielded a temperature of 4100°K for the effective vibrational temperature.

SUMMARY

The trails generated by ultra-speed pellets exhibit several types of emission spectra. These types are: continuous, line, and band. Furthermore, there is a marked difference between the spectra that depends on the composition of the pellet, i.e., aluminum is characterized chiefly by AlO bands, magnesium by MgO bands and magnesium lines, magnesium-lithium by lithium lines and MgO bands with a strong continuum. From the intensities of the AlO bands a temperature of 4100°K was calculated.

The time-resolved spectra show that the AlO persists in the excited state twice as long as the aluminum. The vaporized aluminum oxidizes to AlO leaving AlO in an excited state.

[10] G. Herzberg, *Spectra of Diatomic Molecules* (D. Van Nostrand Company, Inc., New York, 1950), second edition, p. 203.

The Reversion Method for Solving Nonlinear Differential Equations

LOUIS A. PIPES*
University of California, Los Angeles, California
(Received July 23, 1951)

The subject considered in this paper is the presentation of a method for the solution of ordinary nonlinear differential equations of a certain type. The method is based on the algebraic procedure used in reverting power series. By the use of this method, ordinary differential equations having certain type of nonlinearity are reduced to systems of linear differential equations. The resulting linear differential equations are then solved by the Laplace transform or operational method. The general procedure is illustrated by means of examples, and the method is applied to nonlinear electric circuits.

I. INTRODUCTION

IT is the purpose of this discussion to present a method that is applicable to the solution of nonlinear differential equations of a certain class. The method under consideration is based on an algebraic procedure that reduces nonlinear differential equations to a system of linear differential equations. When the methods of the Laplace transform theory are applied to the solution of the resulting system of differential equations, the method saves a great deal of numerical effort. This method was suggested to the author by Mr. A. C. Sim, formerly of the British Thomson-Houston Company, Ltd. Other procedures leading to the same general ideas have been published.[1-3]

The author has had occasion to use the reversion method in the study of certain physical problems whose mathematical formulation leads to nonlinear differential equations. It has been found that the method indicates the complexity of the solution before much labor has been expended. In general, the solution of the nonlinear differential equations is expressed in the form of a combination of polynomials and exponential functions.

II. GENERAL DESCRIPTION OF THE METHOD

Let the nonlinear differential equation whose solution is desired be written in the form

$$a_1y+a_2y^2+\cdots+a_5y^5+\cdots=k\phi(t), \tag{2.1}$$

where t is the independent variable; y, the desired dependent variable; k, a constant; $\phi(t)$, a given function of t. The a_i coefficients are, in general, operators, and are functions of the operator, D, where

$$D=d/dt \quad \text{and} \quad a_1\neq 0. \tag{2.2}$$

Assume a solution of the form,

$$y=A_1k+A_2k^2+A_3k^3+\cdots. \tag{2.3}$$

The A_i coefficients may be determined by substituting (2.3) into (2.1) and equating coefficients of equal powers of k. If this is done, the following equations for the coefficients, A_1, A_2, A_3, etc. are obtained.

$$A_1=\phi(t)/a_1. \tag{2.4}$$

$$A_2=-a_2A_1^2/a_1. \tag{2.5}$$

$$A_3=-(1/a_1)\{2a_2A_1A_2+a_3A_1^3\}. \tag{2.6}$$

$$A_4=-(1/a_1)\{a_2(A_2^2+2A_1A_3)+3a_3A_1^2A_2+a_4A_1^4\}. \tag{2.7}$$

$$A_5=-(1/a_1)\{2a_2(A_1A_4+A_2A_3)+3a_3(A_1A_2^2+A_1^2A_3)+4a_4A_1^3A_2+a_5A_1^5\}. \tag{2.8}$$

$$A_6=-(1/a_1)\{a_2(A_3^2+2A_1A_5+2A_2A_4)+a_3(A_2^3+3A_1^2A_4+6A_1A_2A_3)+2a_4(2A_1^2A_2^2+2A_1^3A_3)+5a_5A_1^4A_2+a_6A_1^6\}. \tag{2.9}$$

$$A_7=-(1/a_1)\{2a_2(A_1A_6+A_2A_5+A_3A_4)+3a_3(A_1^2A_5+A_2^2A_3+A_1A_3^2+2A_1A_2A_4)+4a_4(A_1A_2^3+A_1^3A_4+3A_1^2A_2A_3)+5a_5(2A_1^3A_2^2+A_1^4A_3)+6a_6A_1^5A_2+a_7A_1^7\}. \tag{2.10}$$

Additional coefficients may be found in Van Orstrand's paper,[4] in which a list of the first thirteen coefficients is given.

III. EXAMPLES ILLUSTRATING THE METHOD

The general procedure of the method of reversion may best be illustrated by applying it to several special cases.

Example I

Consider the differential equations

$$dy/dt-y^2=1 \quad \text{with the initial condition that} \quad y=0 \text{ at } t=0. \tag{1}$$

If this is compared with the general equation, (2.1), it is seen that we must have

$$a_1=d/dt=D, \tag{2}$$

and

$$a_2=-1, \quad k\phi(t)=1, \quad k=1. \tag{3}$$

* Professor of Engineering.
[1] L. A. Pipes, J. Acoust. Soc. Am. **10**, 29 (July, 1938).
[2] L. A. Pipes, J. Appl. Phys. **13**, 117 (1942).
[3] C. S. Roys, Proc. Nat. Electron. Conf. p. 663 (1947).
[4] C. E. Van Orstrand, Phil. Mag. **19**, 366 (1910).

Hence from (2.4) we see that

$$A_1=1/a_1, \quad \text{or} \quad DA_1=1. \tag{4}$$

This is a differential equation that determines the function A_1.

Now

$$dA_1/dt=1 \quad \text{or} \quad A_1=t+\text{constant}. \tag{5}$$

In the method of reversion, the initial conditions are impressed on the function A_1 as if it were the entire solution. Hence the initial condition $y=0$ at $t=0$ leads to

$$A_1=t. \tag{6}$$

The function A_2 is determined from Eq. (2.5) and is given by

$$A_2=-a_2A_1^2/a_1, \tag{7}$$

or

$$dA_2/dt=t^2. \tag{8}$$

This integrates into

$$A_2=t^3/3. \tag{9}$$

The constant of integration is zero because the subsequent functions $A_2, A_3, \cdots$ etc. are determined by the condition that they have zero initial values. Equation (2.6) enables the function A_3 to be determined. This is given by

$$A_3=-1/a_1(-2t^4/3), \tag{10}$$

or

$$dA_3/dt=2t^5/3. \tag{11}$$

Hence

$$A_3=2t^5/15. \tag{12}$$

Similarly, the other terms, $A_4, A_5, \cdots$ may be determined. The solution is given by (2.3) and in this case the result

$$y=t+t^3/3+2t^5/15+\cdots \tag{13}$$

is obtained.

The original differential equations, 1, may be written in the form

$$dy/dt=1+y^2 \quad \text{or} \quad dy/(1+y^2)=dt. \tag{14}$$

Hence

$$\tan-1y=t, \tag{15}$$

or

$$y=\tan t, \quad \text{when } y=0 \text{ at } t=0. \tag{16}$$

The series (13) may be recognized as the Taylor's series expansion of the function $\tan t$.

Example II

Let it be required to solve the following differential equation.

$$(dy/dt)+\alpha y^2=t. \tag{1}$$

with the initial condition that $y=y_0$ at $t=0$. The reversion procedure is simplified if the following change of variable is made.

$$v=y-y_0 \quad \text{or} \quad y=v+y_0. \tag{2}$$

This simplifies the initial condition in the variable v giving $v=0$ at $t=0$. This change in variable transforms Eq. (1) into

$$(dv/dt)+\alpha(v^2+2vy_0+y_0^2)=t, \tag{3}$$

or

$$(dv/dt)+2y_0\alpha v+\alpha v^2=t-\alpha y_0^2. \tag{4}$$

If this is compared with (2.1) it is seen that;

$$a_1=[(d/dt)+2y_0\alpha], \tag{5}$$

$$a_2=\alpha, \tag{6}$$

$$k=1, \tag{7}$$

and

$$\phi(t)=(t-\alpha y_0^2). \tag{8}$$

From (2.4) the differential equation for A_1 is obtained in the form

$$a_1A_1=t-\alpha y_0^2, \tag{9}$$

or

$$[(d/dt)+2y_0\alpha]A_1=t-\alpha y_0^2. \tag{10}$$

Since the initial conditions on v are that $v=0$ at $t=0$, it is necessary that $A_1=0$ at $t=0$.

Equation (10) may most easily be solved by the use of the method of the Laplace transform.[5]

If

$$LA_1=\bar{A}_1, \tag{11}$$

and the Laplace transform of both members of the equation (10) is taken, the equation

$$(p+2y_0\alpha)\bar{A}_1=(1/p)-\alpha y_0^2 \tag{12}$$

is obtained.

Hence

$$\bar{A}_1=\frac{1}{p(p+2y_0\alpha)}-\frac{\alpha y_0^2}{(p+2y_0\alpha)}. \tag{13}$$

The inverse transform of this equation is

$$A_1=\frac{t}{2y_0\alpha}-\frac{1}{4y_0^2\alpha^2}+\frac{\epsilon^{-2y_0\alpha t}}{4y_0^2\alpha^2}-\alpha y_0^2\left(\frac{1-\epsilon^{-2y_0\alpha t}}{2y_0\alpha}\right). \tag{14}$$

The function A_2 is now determined from (2.5) and satisfies the differential equation

$$\left(\frac{d}{dt}+2y_0\alpha\right)A_2=-\alpha A_1^2$$

$$=-\alpha\left\{\frac{t}{2y_0\alpha}+\frac{1}{4y_0^2\alpha^2}(\epsilon-1^{-2y_0\alpha t})\right.$$

$$\left.-\frac{y_0}{2}(1-\epsilon^{-2y_0\alpha t})\right\}^2. \tag{15}$$

[5] See, for example, Louis A. Pipes, *Applied Mathematics for Engineers and Physicists* (McGraw-Hill Book Company, Inc., New York, 1946), Chapter VI.

This equation may be solved by the Laplace transform method although the procedure is somewhat tedious. The higher order terms A_3, A_4, etc., may be obtained in a similar manner.

The solution is then given by

$$v=A_1+A_2+A_3+\cdots, \tag{16}$$

or

$$y=y_0+A_1+A_2+A_3+\cdots. \tag{17}$$

Example III

The equation

$$(dy/dt)+\alpha y^2=t \tag{1}$$

discussed previously may be solved in another form without introducing a change in variable. The equation as it stands gives the following values:

$$a_1=d/dt, \tag{2}$$

$$a_2=\alpha, \tag{3}$$

$$k=1, \tag{4}$$

$$\phi(t)=t. \tag{5}$$

Hence the equation for A_1 is

$$dA_1/dt=t. \tag{6}$$

The solution of this equation subject to $A_1=y_0$ at $t=0$ is

$$A_1=(t^2/2)+y_0. \tag{7}$$

The equation for A_2 is

$$\frac{dA_2}{dt}=-\alpha\left(\frac{t^2}{2}+y_0\right)^2, \tag{8}$$

or

$$A_2=-\alpha\left(\frac{t^5}{20}+\frac{y_0t^3}{3}+y_0^2t\right), \tag{9}$$

where use has been made of the condition that $y=y_0$ at $t=0$. The higher order terms may be easily computed and lead to polynomials in t. The solution is of the form

$$y=\left(\frac{t^2}{2}+y_0\right)-\alpha\left(\frac{t^5}{20}+\frac{y_0t^3}{3}+y_0^2t\right)+\cdots, \tag{10}$$

and has a different form from the one obtained in Example II above.

Example IV

The differential equation of an anharmonic oscillator is:

$$(d^2y/dt^2)+\omega_0^2y+\alpha y^2=B\cos\omega t, \tag{1}$$

where $\omega_0\neq\omega$.

Let it be required to solve this differential equation subject to the condition that $y=\dot{y}=0$ at $t=0$.

In this case we have

$$a_1=[(d^2/dt^2)+\omega_0^2], \tag{2}$$

$$a_2=\alpha, \tag{3}$$

$$k=1, \tag{4}$$

$$\phi(t)=B\cos\omega t. \tag{5}$$

The differential equation for A_1 is, in this case,

$$[(d^2/dt^2)+\omega_0^2]A_1=B\cos\omega t. \tag{6}$$

If the Laplace transform of both sides of this equation is taken, we obtain

$$(p^2+\omega_0^2)\bar{A}_1=Bp^2/p^2+\omega^2, \tag{7}$$

where $\bar{A}_1$ is the Laplace transform of A_1, and use has been made of the fact that $A_1=\dot{A}_1=0$ at $t=0$.

Hence we have

$$\bar{A}_1=B\frac{p^2}{(p^2+\omega_0^2)(p^2+\omega^2)}, \tag{8}$$

and the inverse Laplace transform from a table of transforms gives

$$A_1=\frac{B}{(\omega^2-\omega_0^2)}[\cos\omega_0t-\cos\omega t], \tag{9}$$

under the hypothesis that $\omega_0\neq\omega$.

The differential equation for A_2 is

$$\left[\frac{d^2}{dt^2}+\omega_0^2\right]A_2=-\alpha A_1^2. \tag{10}$$

Now

$$A_1^2=\frac{B^2}{(\omega^2-\omega_0^2)^2}\{\cos^2\omega_0t+\cos^2\omega t-2\cos\omega_0t\cos\omega t\}. \tag{11}$$

In order to compute the Laplace transform of A_1^2 and to save writing, let us write

$$T(\omega)=1/(p^2+\omega^2). \tag{12}$$

We then have;

$$L\cos^2\omega_0t=\tfrac{1}{2}[1+p^2T(2\omega_0)]; \tag{13}$$

$$L\cos^2\omega t=\tfrac{1}{2}[1+p^2T(2\omega)]; \tag{14}$$

$$L\cos\omega_0t\cos\omega t=(p^2/2)[T(\omega_0+\omega)+T(\omega_0-\omega)]. \tag{15}$$

Hence on taking the Laplace transform of Eq. (10), we obtain;

$$\bar{A}_2=-\frac{\alpha}{(p^2+\omega_0^2)}\frac{B^2}{(\omega^2-\omega_0^2)}2\left\{1+\frac{p^2}{2}T(2\omega_0)+\frac{p^2}{2}T(2\omega)\right.$$

$$\left.-p^2T(\omega_0+\omega)-P_1^2T(\omega_0-\omega)\right\}, \tag{16}$$

or

$$\bar{A}_2=\frac{-\alpha B^2}{(\omega^2-\omega_0^2)^2}\left\{T(\omega_0)+\frac{p^2}{2}T(\omega_0)T(2\omega_0)+\frac{p^2}{2}T(\omega_0)T(2\omega)\right.$$

$$\left.-p^2T(\omega_0)T(\omega_0+\omega)-p^2T(\omega_0)T(\omega_0-\omega)\right\}, \quad (17)$$

To compute the inverse transform of $\bar{A}_2$ use may be made of the equations

$$L-1p^2T(a)T(b)=\frac{\cos at-\cos bt}{(b^2-a^2)}; \quad (18)$$

$$L-1T(a)=\frac{1-\cos at}{a^2}. \quad (19)$$

Hence, on taking the inverse transform of $\bar{A}_2$, we obtain:

$$A_2=\frac{-\alpha B^2}{(\omega^2-\omega_0^2)}2\left\{\frac{1-\cos\omega_0 t}{\omega_0^2}+\frac{1}{2}\frac{(\cos\omega_0 t-\cos 2\omega_0 t)}{3\omega_0^2}\right.$$

$$+\frac{1}{2}\frac{(\cos\omega_0 t-\cos 2\omega t)}{4\omega^2-\omega_0^2}-\frac{\cos\omega_0 t-\cos(\omega_0+\omega)t}{\omega(2\omega_0+\omega)}$$

$$\left.-\frac{\cos\omega_0 t-\cos(\omega_0-\omega)t}{\omega(\omega-2\omega_0)}\right\}. \quad (20)$$

The solution of (1) is therefore given by

$$y=A_1(t)+A_2(t)+\cdots. \quad (21)$$

Equation (1) was first proposed by Helmholtz to explain the theory of combination tones of music.

Example V

Series Electric Circuit with Nonlinear Inductor

An example of some practical importance may be obtained from electrical circuit theory.

Consider the electrical circuit illustrated in Fig. 1.

The differential equation that governs the flow of current in the electric circuit of Fig. 1 is

$$(Ldi/dt)+Ri+(d\phi/dt)=E \quad t>0, \quad (1)$$

where the switch is supposed closed at $t=0$ so that $i=0$ at $t=0$.

ϕ is the total flux linkage of the nonlinear inductor and is in general a complicated function of i, depending on the nature of the $B-H$ curve of the material of the inductor. A typical analytical expression that is sometimes assumed for the functional relation between ϕ and i is

$$\phi=\tanh i=i-(i^3/3)+(2i^5/15)-\cdots, \quad (2)$$

or

$$\phi=\alpha_0 i-\alpha i^3. \quad (3)$$

Let ϕ be represented by (3). Then the Eq. (1) may be written in the form

$$(Ldi/dt)+Ri+(d/dt)(\alpha_0 i-\alpha i^3)=E, \quad (4)$$

or in terms of the operator $D=d/dt$, we have

$$(L+\alpha_0)Di+Ri-\alpha Di^3=E, \quad (5)$$

or

$$Di+\left(\frac{R}{L+\alpha_0}\right)i-\left(\frac{\alpha}{L+\alpha_0}\right)Di^3=\frac{E}{(L+\alpha_0)}. \quad (6)$$

Now let

$$b=R/(L+\alpha_0). \quad (7)$$

Hence (6) may be written in the form

$$Di+bi-(b\alpha Di^3/R)=Eb/R. \quad (8)$$

In this equation, we have

$$a_1=(D+b), \quad (9)$$

$$a_2=0, \quad (10)$$

$$a_3=-b\alpha D/R, \quad (11)$$

$$k=1, \quad (12)$$

$$\phi(t)=Eb/R. \quad (13)$$

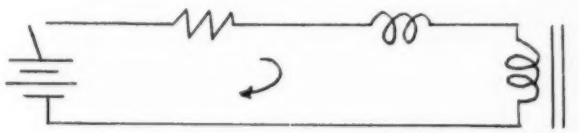


FIG. 1. Ferromagnetic material.

The differential equation for A_1 is

$$[(d/dt)+b]A_1=Eb/R, \quad (14)$$

with the initial condition that $A_1=0$ at $t=0$. The Laplace transform of this equation is

$$\bar{A}_1=Eb/R(p+b). \quad (15)$$

The inverse Laplace transform is

$$A_1=E/R(1-\epsilon^{-bt}). \quad (16)$$

The function A_2 if given by

$$A_2=-(a_2/a_1)A_1^2=0 \quad \text{since} \quad a_2=0. \quad (17)$$

The function A_3 is given by

$$A_3=-a_3A_1^3/a_1, \quad (18)$$

hence, the differential equation for A_3 is given by

$$[(d/dt)+b]A_3=b\alpha dA_1^3/Rdt, \quad (19)$$

$$A_1^3=E^3/R^3(1-\epsilon^{-bt})^3$$
$$=E^3/R^3(1-3\epsilon^{-bt}+3\epsilon^{-2bt}-\epsilon^{-3bt}). \quad (20)$$

If this is substituted into (19) and the Laplace transform of both members of the equation is taken, we

obtain

$$(p+b)\bar{A}_3=\frac{b\alpha}{R}\frac{E^3}{R^3}\left(1-\frac{3p}{p+b}+\frac{3p}{p+2b}-\frac{p}{p+3b}\right). \quad (21)$$

Hence,

$$\bar{A}_3=\frac{b\alpha E^3}{R^4}\left\{\frac{p}{p+b}-\frac{3p^2}{(p+b)^2}+\frac{3p^2}{(p+b)(p+2b)}-\frac{p^2}{(p+b)(p+3b)}\right\}. \quad (22)$$

The inverse Laplace transform of $\bar{A}_3$ may be easily obtained from a table of Laplace transforms, and the result is

$$A_3=\frac{b\alpha E^3}{R^4}\left\{\epsilon^{-bt}-3\epsilon^{-bt}(1-bt)+3\frac{(b\epsilon^{-bt}-2b\epsilon^{-2bt})}{-b}-\frac{(b\epsilon^{-bt}-3b\epsilon^{-3bt})}{-2b}\right\}. \quad (23)$$

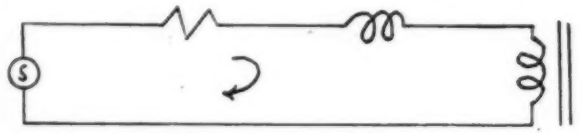


FIG. 2. Ferromagnetic core.

Hence, the current in the circuit is given by the equation:

$$i=A_1+A_3+\cdots,$$

$$=\frac{E}{R}(1-\epsilon^{-bt})+\frac{b\alpha E^3}{R^4}\left\{\epsilon^{-bt}-3\epsilon^{-bt}(1-bt)-3(\epsilon^{-bt}-2\epsilon^{-2bt})+\frac{\epsilon^{-bt}-3\epsilon^{-bt}}{2}\right\}+\cdots. \quad (24)$$

Example VI

Periodic Electromotive Force Applied to a Nonlinear Inductor

Consider the electric circuit of Fig. 2.

Let the periodic electromotive force be applied at $t=0$. The differential equation satisfied by the current for $t>0$ is:

$$Ldi/dt+Ri+d\phi/dt=E\sin wt. \quad (1)$$

If the nonlinear relation between the flux, ϕ, and the current is of the ferromagnetic core, it is represented by the idealized relation,

$$\phi=c_1i-c_3i^3, \quad (2)$$

where c_1 and c_3 are constants determined from the saturation curve of the ferromagnetic material of the core, then Eq. (1) may be written in the form

$$Ldi/dt+Ri+(d/dt)(c_1i+c_3i^3)=E\sin wt. \quad (3)$$

Let the notation

$$b=R/(L+c_1) \quad (4)$$

be introduced. Equation (3) may be written in the form

$$Di+bi-bc_3Di^3/R=Eb\sin wt/R, \quad (5)$$

where $D=d/dt$.

If this equation is compared with (2.1) it is seen that

$$a_1=D+b, \quad a_2=0, \quad a_3=-bc_3D/R, \quad (6)$$

and all the other a's are zero.

In this case, Eq. (2.4) is

$$a_1A_1=(D+b)A_1=Eb\cos wt/R. \quad (7)$$

To solve the differential equation (7) for A_1, let the Laplace transform of A_1, $\bar{A}_1$, be introduced and the Laplace transforms of both members of (7) be taken. The transformed equation is

$$(p+b)\bar{A}_1=\frac{Eb}{R}\frac{wp}{(p^2+w^2)}, \quad (8)$$

where use has been made of the initial condition that

$$A'(0)=A(0)=0. \quad (9)$$

Hence $\bar{A}_1$ is given by

$$\bar{A}_1=\frac{wEb}{R}\frac{p}{(p+b)(p^2+w^2)}. \quad (10)$$

From a table of Laplace transforms the following inverse transform is obtained,

$$A_1=\frac{Eb}{R}\left[\frac{we^{-bt}}{Z^2}+\frac{\sin(wt-\theta)}{Z}\right], \quad (11)$$

where

$$Z=[(b^2+w^2)]^{\frac{1}{2}} \quad \text{and} \quad \theta=\tan^{-1}(w/b). \quad (12)$$

To continue the solution, it is now necessary to compute the sequence of functions, A_2, A_3, etc. Equation (2.5) reduces in this case to

$$A_2=-a_2A_1^2/a_1=0; \quad \text{since } a_2=0, \quad (13)$$

A_3 is given by (2.6) in the form

$$A_3=-a_3A_1^3/a_1, \quad (14)$$

or in this case,

$$(D+b)A_3=\frac{Eb^4c_3D}{R^4Z}\left[\frac{we^{-bt}}{Z}+\sin(wt-\theta)\right]^3. \quad (15)$$

In order to solve Eq. (15), let

$$L[we^{-bt}/Z+\sin(wt-\theta)]^3=\phi(p). \quad (16)$$

Then if $\bar{A}_3$ is the transform of A_3, the Laplace transform of the differential equation (15) subject to the

initial conditions $A_3'(0)=A(0)=0$, is

$$\bar{A}_3=\frac{E^3c_3}{Z(L+c_1)^4}\frac{p\phi(p)}{(p+b)}. \tag{17}$$

Now since,

$$L^{-1}\frac{p}{(p+b)}=e^{-bt}, \tag{18}$$

then the Faltung theorem of the Laplace transform theory leads to the following inverse transform of (17),

$$A_3=\frac{E^3c_3}{Z(L+c_1)^4}\frac{d}{dt}\int_0^t e^{-b(t-u)} \times\left[\frac{we^{-bu}}{Z}+\sin(wu-\theta)\right]^3 du. \tag{19}$$

If the next approximation of the solution is desired, it can be seen from (2.7) and (2.8) that,

$$A_4=0, \quad a_1A_5=-3a_3A_1^2A_3. \tag{20}$$

The current in the circuit is therefore given by

$$i=A_1+A_3+A_5+\cdots,$$

$$=\frac{E}{Z(L+c_1)}\left[\frac{we^{-bt}}{Z}+\sin(wt-\theta)\right]+\frac{E^3c_3}{Z(L+c_1)^4}\frac{d}{dt} \times\int_0^t e^{-b(t-u)}\left[\frac{we^{-bu}}{Z}+\sin(wu-\theta)\right]^3 du. \tag{21}$$

Example VII

General Circuit in Series with a Nonlinear Inductor

The analytical procedure described above may be applied to the case of a general circuit in series with a nonlinear inductor as shown in Fig. 3.

If the circuit of Fig. 3 has an operational impedance $Z(D)$, the differential equation for the current is

$$Z(D)i+(d\phi/dt)=e(t), \quad D=d/dt, \tag{1}$$

where $e(t)$ is the applied electromotive force a given function of time. Let it be assumed to the switch is closed at $t=0$ so that $i=0$ at $t=0$. In general, the flux, ϕ, is an odd function of the current, i, that for analytical purposes may be approximated by a series of the following form,

$$\phi=c_1i+c_3i^3+c_5i^5+\cdots,$$
$$=\sum_{n\,\text{odd}} c_ni^n \quad \text{where the } c\text{'s are constants.} \tag{2}$$

If this is substituted into (1), we have,

$$Z(D)i+\sum c_nDi^n=e(t), \tag{3}$$

or

$$Z(D)i+c_1Di+c_nDi^n=e(t). \tag{4}$$

Hence, comparing this with Eq. (2.1) we obtain

$$a_1(D)=Z(D)+c_1D \tag{5}$$

$$a_n=\begin{cases} c_n & n=3,5,7,\cdots \\ 0 & n=2,4,6,8,\cdots. \end{cases} \tag{6}$$

In this case, Eq. (2.4) may be written in the form,

$$[Z(D)+c_1D]A_1=e(t). \tag{7}$$

If $\bar{e}(p)$ is the Laplace transform of the electromotive force, $e(t)$, and $\bar{A}(p)$ the transform of $A(t)$, then the Laplace transform of the differential equation (7) is

$$\bar{A}_1(p)=\frac{\bar{e}(p)}{Z(p)+c_1p}=\frac{\bar{e}(p)}{a_1(p)}. \tag{8}$$

Hence,

$$A_1(t)=L^{-1}\bar{A}(p). \tag{9}$$

Since $a_2=0$, $A_2=0$. The Laplace transform of A_3 is given by

$$\bar{A}_3=-a_3(p)/a_1(p)=-\frac{c_3pLA_1^3}{Z(p)+c_1p}. \tag{10}$$

$A_4=0$ because a_2 and a_4 are equal to zero. The Laplace transform of A_5 is given by

$$\bar{A}_5=-\frac{1}{a_1(p)}[3a_3(p)LA_1^2LA_3+a_5(p)LA_1^5]. \tag{11}$$

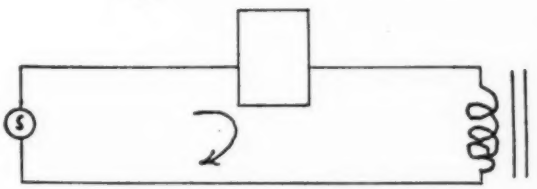


FIG. 3. Ferromagnetic core.

Further terms may be obtained in a similar manner, and the current in the circuit is given by

$$i=A_1+A_3+A_5+\cdots. \tag{12}$$

CONCLUSION

From the preceding examples, the power of the method of reversion can be appreciated. This method is closely associated with the method of Picard for solving differential equations. A study of the convergence of the process indicates that convergence is rapid if the nonlinearities involved are small, and in such cases it is only necessary to compute a few terms of the series of A's to obtain an accurate solution of the differential equation under consideration. The economy of labor effected by the use of the reversion method may be seen by comparing it with other methods of nonlinear mechanics.[6]

[6] See, for example, E. G. Keller, *Mathematics of Modern Engineering* (John Wiley and Sons, Inc., New York, 1942), Vol. II, pp. 294–301, where problems similar to those given previously are treated.

Space Charge and Transit Time Considerations in Planar Diodes for Relativistic Velocities

HENRY F. IVEY
Research Department, Westinghouse Electric Corporation, Bloomfield, New Jersey

(Received August 8, 1951)

The usual Child-Langmuir equation is extended to the case of relativistic velocities for a diode with parallel plane electrodes. The solutions obtained are valid for any value of the accelerating voltage. As the variation of mass with velocity becomes important, the exponent giving the dependence of current density on anode voltage becomes less than three-halves, and for very large values of accelerating potential approaches unity. The variation with anode voltage of the transit time, both for the space-charge-free case and for the space-charge-limited case, has been calculated. The potential distribution across the diode is also discussed.

1. INTRODUCTION

THE usual Child-Langmuir equation[1,2] for space-charge-limited current flow in a planar diode,

$$J=(\sqrt{2}/9\pi)(e/m_0)^{\frac{1}{2}}(V^{\frac{3}{2}}/d^2), \tag{1}$$

is derived under the assumption that the relativistic variation of the mass of the charge carriers with velocity can be neglected. It is of interest to extend this equation to the relativistic case. It is obvious that the increase in mass at the higher velocities will result in a smaller current density than that indicated by Eq. (1). Harmon[3] states that to a first approximation the result is

$$J\propto V^{\frac{3}{2}}\left(1-(3/28)\frac{eV}{m_0c^2}+\cdots\right), \tag{2}$$

but the derivation is not given. Visvanathan[4] has recently studied this problem, but his work unfortunately includes serious mathematical errors. If these are corrected it can be shown that the result is the same as that of Harmon. The method used by Visvanathan cannot very conveniently be extended to obtain further approximations, however.*

Solutions of this problem which are applicable to any value of accelerating voltage are given below. The transit time of the charge carriers, in both the space-charge-limited and the space-charge-free cases, and the potential distribution across the diode have also been worked out. The considerations apply only to diodes with plane parallel electrodes and negligible initial velocities of emission. Because of the high voltages involved, however, this last restriction is actually of no practical importance. The numerical constants have here been evaluated for electronic flow, but the principles are obviously applicable to particles of any mass. For the acceleration of heavier particles, much higher potentials are naturally required to attain relativistic velocities.

2. THE FUNDAMENTAL EQUATIONS

In any space-charge problem the three fundamental relations to be satisfied are Poisson's equation, which for this simple one-dimensional case becomes

$$d^2V/dx^2=4\pi\rho, \tag{3}$$

where ρ is the (negative) space charge density, the equation of continuity,

$$J=\rho v, \tag{4}$$

and the conservation of energy equation. In this case the energy equation is

$$eV=mc^2-m_0c^2=m_0c^2\left(\frac{1}{(1-\beta^2)^{\frac{1}{2}}}-1\right), \tag{5}$$

where, as usual, β signifies the ratio of the particle velocity to the velocity of light, or $\beta=v/c$. It is convenient also to introduce the quantity α which is defined as the ratio of the kinetic energy of the particle to the rest-mass energy ($m_0c^2=0.511$ Mev),

$$\alpha=eV/m_0c^2. \tag{6}$$

From Eq. (5) a relation between α and β may easily be obtained with the result

$$\beta=[\alpha(\alpha+2)]^{\frac{1}{2}}/(\alpha+1). \tag{7}$$

Equations (3), (4), and (7) can be combined to give a differential equation for α,

$$(d^2\alpha/dx^2)=(e/m_0c^2)\cdot(4\pi J/c\beta)$$
$$=(4\pi eJ/m_0c^3)\cdot((\alpha+1)/[\alpha(\alpha+2)]^{\frac{1}{2}}). \tag{8}$$

This may be integrated once by introducing the variable $y=d\alpha/dx$, so that

$$(d^2\alpha/dx^2)=(dy/dx)=(d\alpha/dx)\cdot(dy/d\alpha)=y(dy/d\alpha), \tag{9}$$

[1] C. D. Child, Phys. Rev. **32**, 492 (1911).
[2] I. Langmuir, Phys. Rev. **2**, 450 (1913).
[3] W. W. Harmon, Proc. Inst. Radio Engrs. **37**, 1038 (1949).
[4] S. Visvanathan, Can. J. Phys. **29**, 159 (1951).
* It has recently come to the author's attention that C. C. Wang (Proc. Inst. Radio Engrs. **38**, 135 (1950)) in considering another problem has also given a series solution for the planar diode which applies to Region I in the nomenclature of the present paper.

and

$$y^2=\frac{8\pi eJ}{m_0c^3}\int\frac{(\alpha+1)d\alpha}{[\alpha(\alpha+2)]^{\frac{1}{2}}}=\frac{8\pi eJ}{m_0c^3}[\alpha(\alpha+2)]^{\frac{1}{2}}+C. \quad (10)$$

For complete space charge limitation the electric field (dV/dx) is zero at the cathode $(x=0)$, which here is also taken as the reference point for potential $(V=0)$. The constant of integration thus vanishes and

$$y=(d\alpha/dx)=(8\pi eJ/m_0c^3)^{\frac{1}{2}}[\alpha(\alpha+2)]^{\frac{1}{2}}, \quad (11)$$

$$x=(m_0c^3/8\pi eJ)^{\frac{1}{2}}F(\alpha), \quad (12)$$

where

$$F(\alpha)=\int_0^\alpha d\alpha/[\alpha(\alpha+2)]^{\frac{1}{2}}. \quad (13)$$

It is obviously convenient to take x equal to the interelectrode spacing, d, and α, accordingly, as that value corresponding to the anode potential. Solving for the current density, J, and introducing J_0, the current density in the nonrelativistic case given by (1) above, or

$$J_0=(\sqrt{2}/9\pi)(m_0c^3/e)(\alpha^{\frac{1}{2}}/d^2), \quad (14)$$

one obtains

$$J=\frac{1}{8\pi}\left(\frac{m_0c^3}{e}\right)\frac{F^2}{d^2}=\frac{9}{8\sqrt{2}}\frac{F^2}{\alpha^{\frac{1}{2}}}J_0. \quad (15)$$

The complete solution for the space-charge-limited current density has thus been reduced to evaluation of the function $F(\alpha)$ given by Eq. (13).

3. THE SPACE-CHARGE-LIMITED CURRENT DENSITY

Evaluation of F is most simply performed by expanding the integrand of Eq. (13) by the binomial theorem and then integrating term by term. Consider first the case for $\alpha<2$ (for convenience this will be called Region I). Expanding in ascending powers of α,

$$[\alpha(\alpha+2)]^{-\frac{1}{2}}=\frac{1}{2^{\frac{1}{2}}}\left[\alpha^{-\frac{1}{2}}-\tfrac{1}{8}\alpha^{\frac{1}{2}}+(5/128)\alpha^{7/4}+\cdots+(-1)^n\frac{(1)(5)\cdots(4n-3)}{n!2^{3n}}\alpha^{(n-\frac{1}{2})}+\cdots\right].$$

$$F_1(\alpha)=(4/3)\frac{\alpha^{\frac{3}{2}}}{2^{\frac{1}{2}}}[1-(3/56)\alpha+(15/1408)\alpha^2-(3/1024)\alpha^3+\cdots]+C_1=(4/3)\frac{\alpha^{\frac{3}{2}}}{2^{\frac{1}{2}}}\gamma_1(\alpha)+C_1, \quad (16)$$

where

$$\gamma_1(\alpha)=\sum_{n=0}^{\infty}a_n\alpha^n, \quad (17)$$

and

$$a_n=(-1)^n3\frac{(1)(5)\cdots(4n-3)}{n!2^{3n}(4n+3)}. \quad (a_0=1) \quad (18)$$

By Eq. (12) $F=0$ for $x=\alpha=0$, so that $C_1=0$ and Eq. (15) becomes

$$J_1/J_0=\gamma_1^2(\alpha). \quad (19)$$

For small values of α, γ_1 approaches unity and J approaches the usual equation,

$$J_0=2.330\times10^{-6}V^{\frac{3}{2}}/d^2 \quad (\text{amp/cm}^2), \quad (20)$$

where V is given in volts.

For $\alpha>2$ (Region II) the integrand in Eq. (13) may be expanded in descending powers of α.

$$[\alpha(\alpha+2)]^{-\frac{1}{2}}=\alpha^{-\frac{1}{2}}-\tfrac{1}{2}\alpha^{-\frac{3}{2}}+\tfrac{5}{8}\alpha^{-5/2}+\cdots+(-1)^n\frac{(1)(5)\cdots(4n-3)}{n!2^n}\alpha^{-(n+\frac{1}{2})}+\cdots$$

$$F_2(\alpha)=2\alpha^{\frac{1}{2}}[1+\tfrac{1}{2}\alpha^{-1}-(5/24)\alpha^{-2}+\tfrac{3}{16}\alpha^{-3}+\cdots]+C_2=2\alpha^{\frac{1}{2}}\gamma_2(\alpha)+C_2, \quad (21)$$

where

$$\gamma_2(\alpha)=\sum_{n=0}^{\infty}b_n\alpha^{-n}+\frac{C_2}{2\alpha^{\frac{1}{2}}}, \quad (22)$$

and

$$b_n=(-1)^{n+1}\frac{(1)(5)\cdots(4n-3)}{n!2^n(2n-1)}. \quad (b_0=1) \quad (23)$$

The constant of integration, C_2, cannot be obtained directly because the point $\alpha=0$ is outside the region of convergence of the expansion employed. This constant will be evaluated below. From Eq. (15)

$$J_2/J_0=(9/2\sqrt{2})(\gamma_2^2(\alpha)/\alpha^{\frac{1}{2}}). \quad (24)$$

As α becomes very large, $\gamma_2(\alpha)$ approaches unity and J approaches

$$J_\infty=(c/2\pi)(V/d^2)=5.309\times10^{-3}V/d^2 \quad (\text{amp/cm}^2). \quad (25)$$

This result may also be obtained quite simply by putting $v=c$ in Poisson's equation and integrating directly. The quantity $\gamma_2^2(\alpha)$ is also obviously equal to J_2/J_∞.

To determine the constant C_2 it is necessary to obtain an expression for $F(\alpha)$ which is convergent in the region around $\alpha=2$, and which will thus overlap Region I and II. Introduce the new variable

$$z^4=\alpha(\alpha+2) \quad \text{or} \quad \alpha=(1+z^4)^{\frac{1}{2}}-1. \quad (26)$$

$F(\alpha)$ now becomes

$$F(z)=2\int(z^2dz/(1+z^4)^{\frac{1}{2}}). \quad (27)$$

The integrand may now be expanded in descending

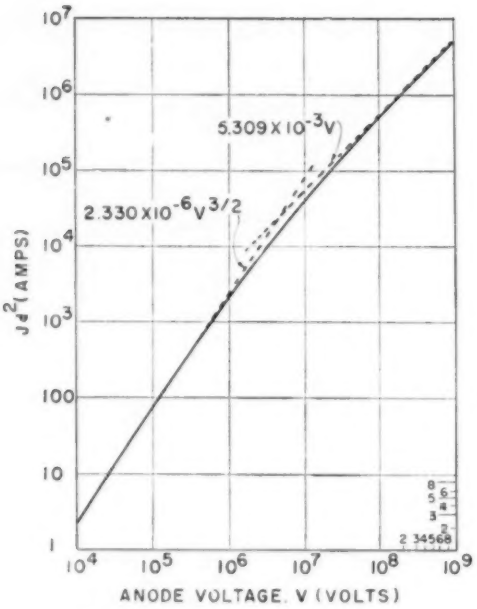


FIG. 1. Space charge characteristics of a planar diode for high anode voltages.

powers of z and an expression valid in the region $z>1$ or $\alpha>0.414$ (Region III) obtained.†

$$z^2(1+z^4)^{-\frac{1}{2}}=1-\tfrac{1}{2}z^{-4}+\tfrac{3}{8}z^{-8}-\tfrac{5}{16}z^{-12}+\cdots$$
$$+(-1)^n\frac{(1)(3)(5)\cdots(2n-1)}{n!2^n}z^{-4n}+\cdots$$
$$F_3(z)=2z(1+\tfrac{1}{6}z^{-4}-(3/56)z^{-8}+(5/176)z^{-12}+\cdots)$$
$$+C_3=2z\sum_{n=0}^{\infty}c_nz^{-4n}+C_3, \quad (28)$$

where

$$c_n=(-1)^{n+1}\frac{(1)(3)\cdots(2n-1)}{n!2^n(4n-1)}. \quad (c_0=1) \quad (29)$$

The constants C_3 and C_2 may be evaluated by equating the appropriate expressions for F at points common to Regions I and III and Regions II and III, respectively. Values of $\alpha=1$ and $\alpha=4$ are convenient for this purpose. The results are

$$C_2=C_3=-1.6931. \quad (30)$$

Results of the above calculations have been plotted in Fig. 1. For small values of α the usual three-halves power law, Eq. (1), holds, but for very high voltages the linear relation (25) is valid. The deviation from the three-halves power law is about 2 percent at 100 kv and about 10 percent at 600 kv. From a practical point of view it must be admitted that the current densities required to achieve space-charge limitation at very high voltages are prohibitive.

Values of the coefficients in the various series solutions employed are given in Table I. All the series converge quite rapidly in the appropriate regions of application. Harmon's approximate solution for small values of α, Eq. (2), follows directly from Eqs. (16) and (19).

4. THE TRANSIT TIME

It is of interest to consider the transit time in both the space-charge-free and the completely space-charge-limited cases. In the former case the variation of potential throughout the diode is linear, so that

$$\alpha/\alpha_a=x/d, \quad (31)$$

where α_a indicates the value of α at the anode ($x=d$). Utilizing Eq. (7) one obtains

$$v=\frac{dx}{dt}=\beta c=c\frac{[\alpha(\alpha+2)]^{\frac{1}{2}}}{(\alpha+1)}, \quad (32)$$

or

$$dt=\frac{d}{c\alpha_a}\frac{(\alpha+1)}{[\alpha(\alpha+2)]^{\frac{1}{2}}}d\alpha.$$

TABLE I. Numerical values of the coefficients in the series for the function $F(\alpha)$.

n	a_n	b_n	c_n
0	+1.0000000	+1.0000000	+1.0000000
1	−0.0535714	+0.5000000	+0.1666667
2	+0.0106534	−0.2083333	−0.0535714
3	−0.0029297	+0.1875000	+0.0284091
4	+0.0009396	−0.2176339	−0.0182292
5	−0.0003299	+0.2877604	+0.0129523
6	+0.0001229	−0.4346591	−0.0098505
7	−0.0000478	+0.6225586	+0.0077582

The transit time is thus

$$T_1=\int_0^{\alpha_a}dt=(d/c\alpha_a)\Big[\alpha(\alpha+2)\Big]^{\frac{1}{2}}_{\alpha=\alpha_a}.$$

Dropping the subscript on α,

$$T_1=d/c[(\alpha+2)/\alpha]^{\frac{1}{2}}. \quad (33)$$

As α becomes very large, T_1 approaches $d/c(=3.336\times10^{-11}\times d$ sec) as is to be expected since in this case the velocity of the electrons is practically constant and equal to c throughout the diode. For very small values of α, the transit time approaches the usual value,[5]

$$T_0=d/c(2/\alpha)^{\frac{1}{2}}=2d(m_0/2eV)^{\frac{1}{2}}$$
$$=3.373\times10^{-8}d/V^{\frac{1}{2}}\text{ sec}), \quad (34)$$

so that

$$T_1/T_0=(\alpha+2/2)^{\frac{1}{2}}. \quad (35)$$

† Equation (27) is of the form employed by Visvanathan[4]. The expressions for Regions I and III are actually sufficient for a complete solution of the problem, but that for Region II (involving α directly) is considered preferable because of the more direct physical interpretation.

[5] K. R. Spangenberg, *Vacuum Tubes* (McGraw-Hill Book Company, Inc., New York, 1948), p. 195.

Turning now to the space-charge-limited case, and utilizing Eqs. (11) and (32), one obtains

$$dt=\left(\frac{m_0c}{8\pi eJ}\right)^{\frac{1}{2}}\frac{(\alpha+1)}{[\alpha(\alpha+2)]^{\frac{1}{2}}}d\alpha.$$
$$T=\int_0^{\alpha}dt=2\left(\frac{m_0c}{8\pi eJ}\right)^{\frac{1}{2}}[\alpha(\alpha+2)]^{\frac{1}{2}}. \tag{36}$$

Substitution of J_0 from Eq. (14) yields

$$T=\frac{3}{2^{\frac{3}{2}}}\frac{d}{c}\left(\frac{J_0}{J}\right)^{\frac{1}{2}}\frac{(\alpha+2)^{\frac{1}{2}}}{\alpha^{\frac{1}{2}}}, \tag{37}$$

or

$$\frac{T}{T_0}=\frac{3}{2}\left(\frac{J_0}{J}\right)^{\frac{1}{2}}\left(\frac{\alpha+2}{2}\right)^{\frac{1}{2}}=\frac{3}{2}\left(\frac{J_0}{J}\cdot\frac{T_1}{T_0}\right)^{\frac{1}{2}}. \tag{38}$$

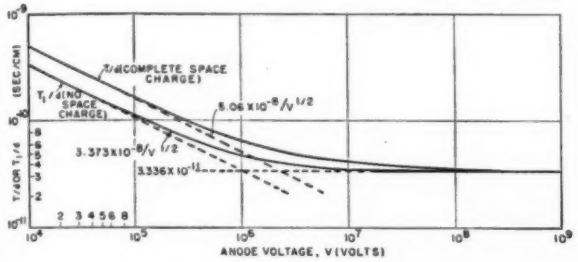


FIG. 2. Transit time in a planar diode as a function of anode voltage.

It is seen that for small values of α the transit time in the presence of space charge is three-halves that in the absence of space charge.[5] For very high values of α the two transit times obviously approach the same limiting value. The variation of the transit time with applied potential in both cases is shown in Fig. 2.

5. THE POTENTIAL DISTRIBUTION

Since the current density must remain constant throughout the diode, it is seen from Eq. (12) that

$$(x/d)=[F(\alpha)/F(\alpha_a)]. \tag{39}$$

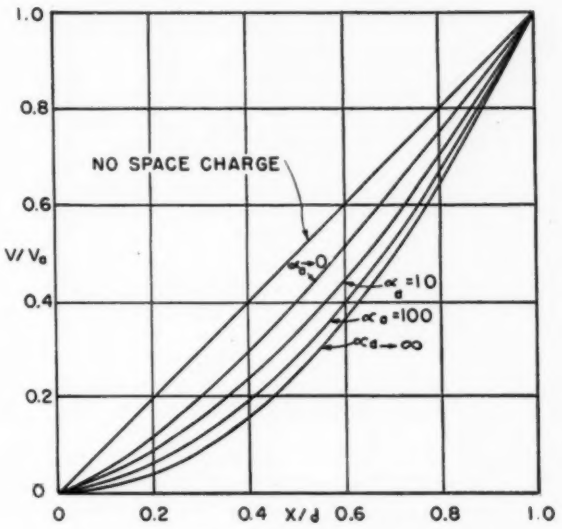


FIG. 3. Potential distribution in a planar diode.

For $\alpha<2$, from Eq. (16),

$$x/d=(\alpha/\alpha_a)^{\frac{4}{3}}[\gamma_1(\alpha)/\gamma_1(\alpha_a)]. \tag{40}$$

For very small values of α_a, γ_1 approaches unity and one obtains the usual nonrelativistic case with the potential varying as the four-thirds power of the distance. The electric field at the anode in this case is four-thirds that in the absence of space charge.

For $\alpha>2$, from Eq. (21),

$$x/d=(\alpha/\alpha_a)^{\frac{1}{2}}[\gamma_2(\alpha)/\gamma_2(\alpha_a)]. \tag{41}$$

For very large values of α_a, γ_2 approaches unit and it is seen that here the potential varies as the square of the distance. The corresponding electric field at the anode is therefore twice that in the absence of space charge or three-halves of that for space-charge limitation at low anode voltages. In Fig. 3 the potential distribution is plotted for the two limiting cases discussed above and also for two intermediate values of α_a.

Mr. R. Gaines Young of this laboratory has kindly checked the mathematics of this paper.

The Structure of Tetragonal Tungsten Trioxide

W. L. KEHL, R. G. HAY, AND D. WAHL*
Gulf Research and Development Company, Pittsburgh, Pennsylvania

(Received July 20, 1951)

A study of tungsten trioxide (WO_3) by differential thermal analysis and x-ray diffraction at temperatures up to 980°C showed a new tetragonal form to exist at temperatures of 740°C and higher. This form has a distorted ReO_3 structure. There are two molecules in a unit cell of dimensions $a_0=5.25$A, $c_0=3.91$A. All atoms lie in the 2- or 4-fold special positions of the space group $P4/nmm$. The tungsten atoms are alternately displaced in opposite directions from the center of charge of the oxygen atoms so as to resemble the antiferroelectric structure proposed by Kittel.

INTRODUCTION

As part of a general study of catalytic oxides the differential thermal curve was obtained for each material studied. This curve provides a record of any temperature differences which occur between the specimen and a standard "inert" sample while both are heated at the same uniform rate, and has been found to be a very sensitive indicator of chemical or structural changes which occur in a material as a result of a change of temperature. The differential thermal curve obtained for tungsten trioxide between 25°C and 1000°C exhibited endothermic peaks with centers at 740°C and 900°C. X-ray diffraction patterns of this material were therefore recorded at 25°C and at several temperatures in the interval from 600°C to 980°C. The pattern taken at 25°C corresponds to the pattern found for WO_3 by Glemser and Sauer,[1] and appears to exhibit the triclinic structure reported[2] for WO_3 at room temperature.

At 740°C and higher the diffraction patterns show that the structure changes to a tetragonal form which the present authors reported earlier[3] and which has also been reported more recently by Ueda and Ichinokawa.[4]

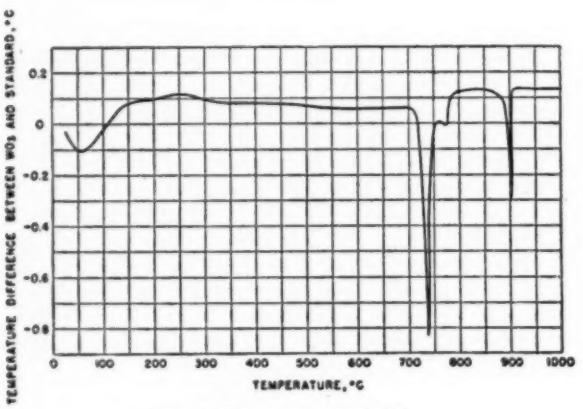


FIG. 1. Thermogram of WO_3.

This form has the same composition as the room temperature form, since the phase change was reversible, as indicated by both the differential thermal curves and the x-ray diffraction patterns at room temperature before and after heating above the transition temperature. The crystal structure and the atomic parameters of this tetragonal form have been determined from measurements on x-ray powder photographs.

EXPERIMENTAL

The apparatus used for the differential thermal analysis consisted of an Inconel sample holder, similar to that described by Berkelhamer,[5] heated in a resistance furnace having an Inconel lining. This lining gives somewhat better heat distribution than a ceramic core. The sample was heated at a rate of approximately 100°C per 15 minutes after reaching 100°C, using a motor-driven variable transformer as a source of power. A positive temperature control was less satisfactory because the off-on cycle tended to be reflected in the differential thermal curve. The signal generated by a chromel-alumel differential thermocouple was amplified by a Leeds & Northrup No. 9835 dc microvolt amplifier and recorded on a ten-millivolt full scale Speedomax. Figure 1 shows the heating part of the differential thermal curve for WO_3. The minima of the peaks occur at 740°C and 900°C, but the curve begins to dip down at somewhat lower temperatures, indicating that the "reaction," or phase change, probably takes place at these lower temperatures. On the cooling part of the curve these peaks appeared at positions about 20°C lower than their positions on the curve in Fig. 1.

The high temperature x-ray diffraction camera was built by the Central Research Laboratories.† This camera has a diameter of 114.6 mm. Provision is made to move the film holder so that five different exposures can be made on one film. A thin-walled, fused-quartz capillary, approximately 0.5 mm in diameter and 37 mm long, serves to support the specimen inside the furnace and camera. The outer end, which contains the specimen, is closed, and the other end is effectively sealed by

* Present address: Sylvania Electric Products Company, Bayside, Long Island.

[1] O. Glemser and H. Sauer, Z. anorg. u. allgem. Chem. **252**, 144 (1943).

[2] H. Breakken, Z. Krist. **78**, 484 (1931).

[3] Kehl, Hay, and Wahl, Bull. Am. Phys. Soc. **26**, No. 2, 18 (3-8-51).

[4] R. Ueda and T. Ichinokawa, Phys. Rev. **82**, 563 (1951).

[5] L. H. Berkelhamer, RI 3762, U. S. Dept. of Int. Bur. of Mines, July, 1944.

† Address: Red Wing, Minnesota.

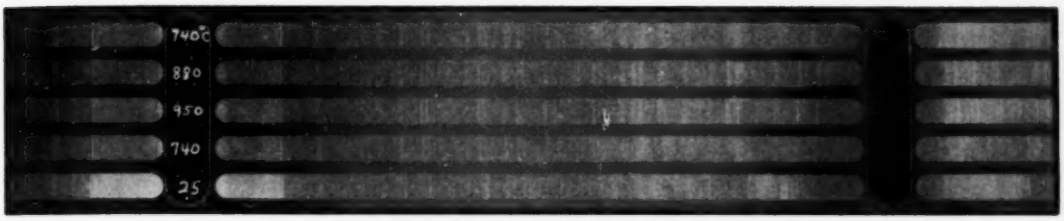


FIG. 2. X-ray diffraction patterns of WO_3 obtained at the temperatures indicated, using filtered CuK radiation.

being firmly embedded in modeling clay which holds the capillary in the supporting spindle. This method is used to protect the specimen from oxidation during heating, since no effort is made to control the atmosphere inside the furnace.

Figure 2 shows a typical film on which are patterns of WO_3 obtained at 740°C, 880°C, 950°C, 740°C, and room temperature, in the order listed. The 2.25A line of platinum, which appears in the pattern obtained at 950°C and the two following patterns, is caused by a thin film of platinum which evaporated from the heater coils and condensed on the outside surface of the fused-quartz capillary during the 950°C exposure.

The material used in this study was commercially available cp tungstic acid anhydride. Later the results were checked by comparison with patterns obtained from a very pure sample‡ of tungstic acid run at room temperature, 770°C and 980°C. This sample was first heated to 800°C in a muffle furnace to drive off the water. The patterns obtained with this material were identical to those obtained earlier from the commercial cp material.

DETERMINATION OF STRUCTURE

In the x-ray diffraction pattern that was obtained at 740°C, the lines in the back-reflection region are broad and poorly resolved so that it is difficult to index them with certainty. At higher temperatures, these back-reflection lines become progressively sharper and are well-defined in the pattern at 950°C. This latter pattern, which represents the same structure as observed between 740°C and 900°C, could be indexed on the basis of a tetragonal lattice, and with it as a guide it was also possible to index the back-reflection lines in the pattern obtained at 770°C. The lattice parameters calculated from the measurements on these photographs were:

$$a_0 = 5.272 \pm 0.002\text{A}, \quad c_0 = 3.920 \pm 0.002\text{A at } 950°\text{C};$$

$$a_0 = 5.250 \pm 0.002\text{A}, \quad c_0 = 3.915 \pm 0.002\text{A at } 770°\text{C}.$$

The only systematic extinctions observed were in the (*hko*) zone when $h+k$ is odd. This is characteristic for space groups $P4/n$ and $P4/nmm$ in the tetragonal system. The unit cell contains two molecules of WO_3, and all atoms must therefore lie in the 2- or 4-fold special positions, which are the same for both space groups. The reported triclinic structure at room temperature is a lower symmetry variation of the ReO_3 cubic structure, and it seemed probable that this tetragonal form would also have a modified ReO_3 type of structure intermediate between the triclinic and ideal cubic forms. Such a structure requires the atoms to occupy the following special positions:

$$2\text{W}: \quad 0\tfrac{1}{2}z_1, \ \tfrac{1}{2}0\bar{z}_1,$$

$$2\text{O}_\text{I}: \quad 0\tfrac{1}{2}z_2, \ \tfrac{1}{2}0\bar{z}_2,$$

$$4\text{O}_\text{II}: \quad \tfrac{1}{4}\tfrac{1}{4}0, \ \tfrac{3}{4}\tfrac{3}{4}0, \ \tfrac{1}{4}\tfrac{3}{4}0, \ \tfrac{3}{4}\tfrac{1}{4}0.$$

It was observed that the intensity of the 002 reflection was weak, which indicated that z_1 could not be zero, that is, that the tungsten atoms were not coplanar, as in the ideal cubic form. In order to determine z_1, the intensities of several of the lower order reflections were calculated and plotted as a function of z_1, as shown in Fig. 3. All calculated intensities were corrected for absorption in the specimen, and a Debye-Waller temperature factor correction corresponding to $B=3\times10^{-16}$ was applied to the tungsten atoms. In addition, the usual multiplicity and Lorentz-polarization corrections for Debye-Scherrer patterns were applied. The best agreement between the intensities so calculated and those observed in the pattern of WO_3 at 950°C was obtained for $z_1=0.06$. On the basis of the modified cubic structure being considered, it was apparent that the O_I atoms must lie on the tetrad axis in the interval between the midpoint of the axis and midway between the

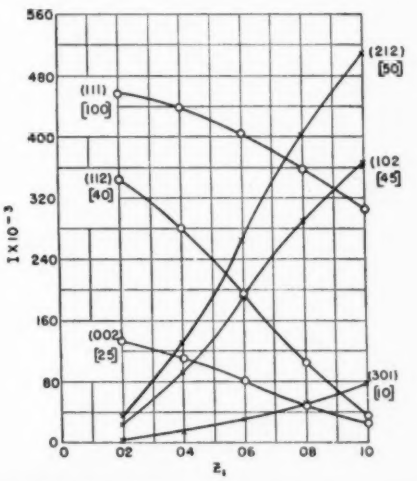


FIG. 3. Plot of calculated intensity *vs* z_1, using $z_2=z_1+\frac{1}{2}$. Numbers in brackets indicate observed relative intensities.

‡ This sample was made and very kindly sent to us by Dr. B. T. Matthias of the Bell Telephone Laboratories.

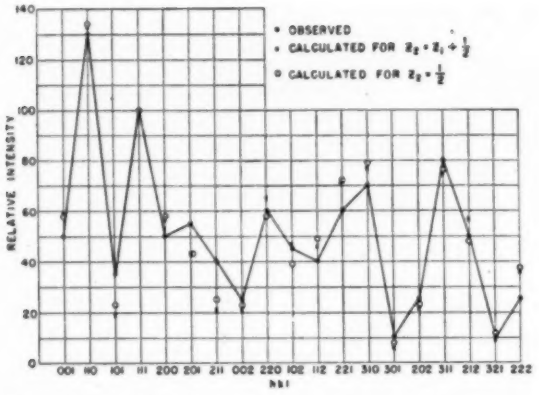


FIG. 4. Comparison of the relative intensities observed and calculated for the tetragonal WO_3 structure.

tungsten atoms, that is between $z_2=\frac{1}{2}$ and $z_2=z_1+\frac{1}{2}$. The relative intensities of the first 19 observed lines were calculated both for assignments $z_1=0.06$, $z_2=\frac{1}{2}$, and $z_1=0.06$, $z_2=z_1+\frac{1}{2}$. The results are shown in Fig. 4, where the values of the observed relative intensities are connected by solid lines. As might be expected, the contribution of the two O_I atoms to the diffracted intensities is generally very small compared to that from the other atoms. Consequently, no decisive evidence was obtained to indicate which value of z_2 was more nearly the correct one. The agreement between the observed and both sets of calculated intensities is considered to be within the limits of error in measuring the intensities. The structure is illustrated in Fig. 5, where (a) is a projection on the basal plane of the oxygens at $z=0$ and the tungstens at $z=\pm 0.06$, and (b) is a projection onto the $x0z$ face.

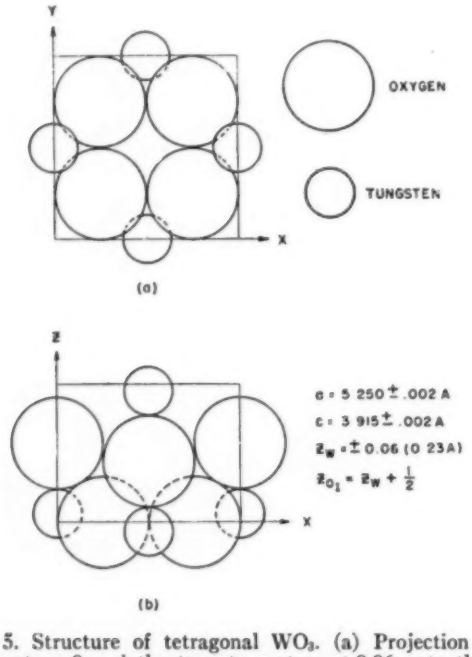


FIG. 5. Structure of tetragonal WO_3. (a) Projection of the oxygens at $z=0$ and the tungstens at $z=\pm 0.06$ onto the basal plane. (b) Projection onto the $x0z$ face.

DISCUSSION

This tetragonal structure, as well as the reported triclinic room-temperature form, may be described as a distorted cubic structure of the ReO_3 type. The idealized cubic form has tungsten atoms at the corners of a cube, each surrounded by six octahedrally coordinated oxygens. Four oxygens lie in a plane which contains the tungsten atoms, and there is one oxygen above and one below this plane for each tungsten atom. Each oxygen is common to two octahedra, which gives the formula WO_3. In the tetragonal form, the tungsten atoms are displaced out of the plane of the four oxygen atoms a distance of 0.06c (approximately 0.23A), as shown in Fig. 5.

The ideal cubic form would contain one molecule per unit cell. The tetragonal form, with two molecules per cell, is related to the cubic cell in the manner shown in Fig. 6. In the lower symmetry triclinic form there are

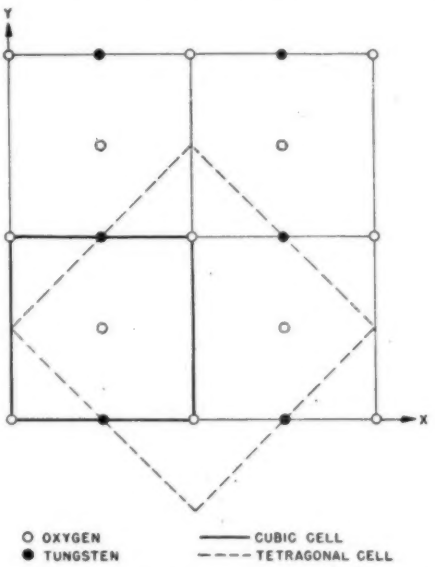


FIG. 6. Relation of tetragonal to the idealized cubic and triclinic unit cells.

four molecules per unit cell, and the a and b axis lengths are each approximately twice the length expected for a in the cubic cell. The c axis length is approximately the same in all three lattices.

Lines in the back-reflection region of the pattern taken at 740°C are broad and rather indistinct, but become increasingly sharper and more distinct in the patterns taken at higher temperatures. Apparently, the disordered structure appearing just above the transition temperature becomes ordered at higher temperatures. It was also noted that the thermal expansion between 770°C and 950°C is 2.33×10^{-5}/C° along a and only 7.10×10^{-6}/C° along c.

A second transition, from tetragonal to the ideal cubic, was expected because of the second peak on the DTA curve at 900°C, but was not observed in the diffraction patterns obtained at temperatures up to

980°C. Consequently, this second peak on the DTA curve is not explained by the expected structural transition to higher symmetry.

Ueda and Ichinokawa[6] have discussed the domain structure of WO_3 and have pointed out that this material exhibits phenomena similar to those of $BaTiO_3$. Matthias[7] has also reported that at room temperature WO_3 behaves like a ferroelectric. More recently Kittel[8] has proposed a theory of antiferroelectric crystals which he defines as crystals in which lines of ions or cells are spontaneously polarized, but with adjacent lines polarized in antiparallel directions. The tetragonal form of WO_3 resembles this proposed structure, since the displacement of each tungsten atom from the plane of the O_{II} atoms results in a dipole. These tungsten atoms are displaced alternately above and below the plane, so that the dipole moments of adjacent dipoles are antiparallel and hence cancel one another. The displacement of 0.23A is of the same order of magnitude as that found by Evans[9] for $BaTiO_3$.

[6] R. Ueda and T. Ichinokawa, Phys. Rev. **80**, 1106 (1950).
[7] B. T. Matthias, Phys. Rev. **76**, 430 (1949).
[8] C. Kittel, Phys. Rev. **82**, 729 (1951). Also private communication.

ACKNOWLEDGMENTS

The authors are greatly indebted to Dr. G. A. Jeffrey, visiting Professor at the University of Pittsburgh, for his valuable suggestions and criticism, and to Dr. Paul D. Foote, Executive Vice-President of the Gulf Research and Development Company, for permission to publish this work.

[9] H. T. Evans, Jr., Acta Cryst. **4**, 377 (1951).

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 2 FEBRUARY, 1952

Space-Charge-Wave Propagation in a Cylindrical Electron Beam of Finite Lateral Extension*

PHILIP PARZEN
Federal Telecommunication Laboratories, Incorporated, Nutley, New Jersey
(Received June 13, 1951)

A method is outlined for obtaining the effect of geometry on the propagation of space-charge waves in a cylindrical electron beam under arbitrary direct current conditions. In particular, the case of the complete-space-charge diode is discussed in detail. The results of this theory are verified by the experiments of Cutler and Quate.

INTRODUCTION

THE nature of the propagation of space-charge waves in cylindrical electron beams of finite lateral extension has been discussed by Hahn[1] and Ramo[2] for the case of an electron beam drifting with a constant velocity v_0 in a tube. According to the latter paper, the field quantities propagate with a wave number Γ given by

$$\Gamma=\frac{\omega}{v_0}\pm\frac{\omega_p}{v_0[1+(T^2/\gamma_0^2)]^{\frac{1}{2}}}, \tag{1}$$

where ω is the angular frequency, $\gamma_0=\omega/v_0$, the plasma frequency

$$\omega_p=(e\rho_0/m\epsilon_0)^{\frac{1}{2}},$$

ρ_0 is the direct-current charge density, the direct-current density $J_0=\rho_0 v_0$, and T/γ_0 is a function of $R=a/b$ and $\gamma_0 b$. T is the radial propagation number in the beam, a is the radius of the drift tube, and b is the radius of the beam.

A plot of T/γ_0 *versus* $\gamma_0 b$ for various values of R is given in Fig. 1. For $b=\infty$ (infinite electron beam), T equals zero, and for $R=1$, $Tb=2.4$. For all other conditions, Tb varies between 0 and 2.4.

Equation (1) may be interpreted physically as follows: The first term denotes the phase shift due to the dc transit angle. The second term denotes the phase shift due to space charge and is a function of geometry. The latter term is proportional to $[J_0/1+(T^2/\gamma_0^2)]^{\frac{1}{2}}$. Thus, it appears that the behavior of this phase shift is governed by an effective direct current density J_{0e}, which is given by

$$J_{0e}=J_0/1+(T^2/\gamma_0^2). \tag{2}$$

Thus, one may obtain the nature of the propagation of space-charge waves in those problems where the dc parameters and geometry may vary along the direction of propagation by replacing J_0 by J_{0e} in the differential equations governing such propagation in an infinite beam. A more rigorous discussion of this replacement process is given in the appendix. This method assumes that any lateral variations of dc parameters are small. The equations for an infinite beam will now be given.

* This work was sponsored by the Signal Corps Engineering Laboratories, Fort Monmouth, New Jersey.
[1] W. C. Hahn, Gen. Elec. Rev. **42**, 258 (1939).
[2] S. Ramo, Proc. Inst. Radio Engrs. **27**, 757 (1939).

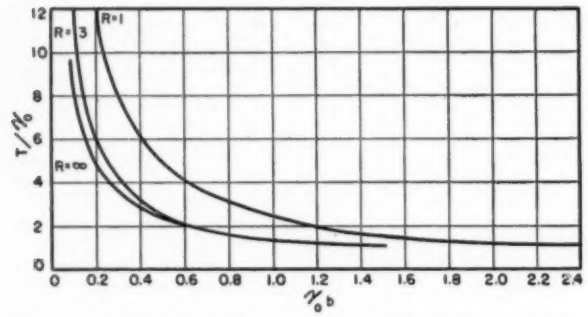


FIG. 1. Correction to plasma frequency in drift tube for beam and drift-tube geometry. R is the (radius of the drift tube)/(radius of the beam), $\gamma_0=\omega/v_{0d}=(2\pi f)/$(beam velocity in the drift tube), b is the beam radius, ω_d is the plasma frequency in the drift tube, $\omega_{d'}$ is the plasma frequency corrected for beam and drift-tube geometry, and $\omega_{d'}/\omega_d=1/[1+(T/\gamma_0)^2]^{\frac{1}{2}}$.

SPACE-CHARGE-WAVE PROPAGATION IN AN INFINITE ELECTRON BEAM

The Maxwell equations for small-signal theory and for a one-dimensional propagating process in an electron beam are

$$\left(j\omega+\frac{\partial v_0}{\partial Z}\right)v+v_0\frac{\partial v_0}{\partial Z}=\frac{e}{m}E, \tag{3}$$

$$\frac{\partial J}{\partial Z}+j\omega\rho=0, \tag{4}$$

$$J+j\omega\epsilon_0 E=0, \tag{5}$$

$$J=\rho_0 v+\rho v_0. \tag{6}$$

Here all field quantities vary as $\exp j\omega t$ and the ac quantity is assumed small with respect to the corresponding dc quantities.

E, J, v, ρ are the ac electric field, current density, velocity, and charge density, respectively, and E_0, J_0, v_0, ρ_0 are the corresponding dc quantities. e is the charge of the electron and is negative, and m is the mass of the electron.

After solving for all field quantities in terms of J, one obtains the following differential equation for J.

$$v_0^3\frac{\partial^2 J}{\partial Z^2}+\left[3v_0^2\frac{\partial v_0}{\partial Z}+j2\omega v_0^2\right]\frac{\partial J}{\partial Z}+\left[\frac{eJ_0}{m\epsilon_0}+j2\omega v_0\frac{\partial v_0}{\partial Z}-\omega^2 v_0\right]J=0. \tag{7}$$

$$v=\frac{v_0}{J_0}\left[J+\frac{v_0}{j\omega}\frac{\partial J}{\partial Z}\right]. \tag{8}$$

Equation (7) can be simplified by replacing Z by the transit time τ. Thus let

$$J=(Y/v_0)\exp-j\omega\tau,$$
$$\tau=\int_0^Z(dZ/v_0). \tag{9}$$

Then Y satisfies the equations

$$(\partial^2 Y/\partial\tau^2)-K^2Y=0, \tag{10}$$

$$K^2=-\frac{1}{v_0}\left[\frac{\partial^2 v_0}{\partial\tau^2}-\frac{eJ_0}{m\epsilon_0}\right], \tag{11}$$

and

$$v=\frac{v_0}{J_0}\left[J+\frac{1}{j\omega}\frac{\partial J}{\partial\tau}\right]. \tag{12}$$

To account for a finite geometry, one may now apply the replacement of J_0 by J_{0e} discussed previously. Thus, Eq. (11) now becomes

$$K^2=-\frac{1}{v_0}\left\{\frac{\partial^2 v_0}{\partial\tau^2}-\frac{eJ_0}{m\epsilon_0}\frac{1}{1+(T^2/\gamma_0^2)}\right\}. \tag{13}$$

Thus, given K as a function of τ, one obtains a solution of Eq. (10) with two arbitrary constants. These constants are then determined by the initial conditions for J and v. Particular consideration will now be given to the case of the complete-space-charge diode.

SPACE-CHARGE WAVES IN A COMPLETE-SPACE-CHARGE DIODE

The case to be considered in Fig. 2 is that of a cylindrical beam of radius b surrounded by a cylinder of radius a in a complete-space-charge diode. Denoting conditions at the cathode and anode by subscripts *1* and *2*, respectively,

$$v_0=v_{02}(Z/d)^{\frac{1}{2}},$$
$$\tau=(3d^{\frac{1}{2}}/v_{02})Z^{\frac{1}{2}}.$$

Hence

$$v_0=\frac{v_{02}^3}{d^2}\left(\frac{\tau}{3}\right)^2. \tag{14}$$

Also

$$v_{02}^3=-\frac{9}{2}\frac{eJ_0}{m\epsilon_0}d^2.$$

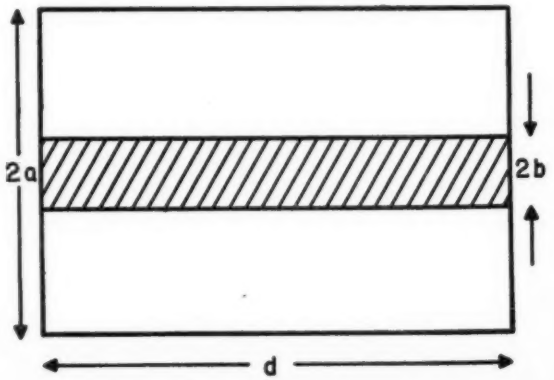


FIG. 2. Complete-space-charge diode.

Now let

$$\phi=g^{\frac{1}{4}}\frac{v_{02}}{d}\frac{\tau}{3};\quad 0<\phi<g^{\frac{1}{4}},$$

$$g=5.76/(\gamma_{02}b)^2,$$

$$G=(Tb/2.4)^2,$$

$$v_0=(v_{02}/g^{\frac{1}{4}})\phi^2. \tag{15}$$

Thus Eq. (10) becomes

$$(\partial^2 Y/\partial\phi^2)-K_\phi^2 Y=0,$$

$$K_\phi^2=2G\phi^2/(1+G\phi^4),$$

$$v=\frac{v_0}{J_0}\left(J+\frac{g^{\frac{1}{4}}}{j3\gamma_{02}d}\frac{\partial J}{\partial\phi}\right). \tag{16}$$

Now applying the WKB method,[3] an approximate solution is

$$Y=C_1P+C_2Q,$$

$$P=\left(\frac{\zeta}{K_\phi}\right)^{\frac{1}{2}}I_{\frac{1}{4}}(\zeta),$$

$$Q=\left(\frac{\zeta}{K_\phi}\right)^{\frac{1}{2}}I_{-\frac{1}{4}}(\zeta),$$

$$\zeta=\int_0^\phi K_\phi d\phi. \tag{17}$$

Here $I_{\frac{1}{4}}$ is the modified Bessel function of order $\frac{1}{4}$. C_1 and C_2 are arbitrary constants to be determined from the initial conditions. The case of $G=1$ for $a=b$ and $G\neq 1$ will now be discussed separately.

(a) $G=1$ for $a=b$

A plot of K_ϕ and ζ *versus* ϕ is given in Fig. 3. From Eqs. (9) and (17)

$$J=\exp\left(-j\frac{3\omega}{g^{\frac{1}{4}}}\frac{d}{v_{02}}\phi\right)\frac{C_1P+C_2Q}{v_0}. \tag{18}$$

For small ϕ

$$K_\phi=1.414\phi\left[1-\frac{1}{2}\phi^4\right],$$

$$\zeta=0.707\phi^2\left[1-\frac{1}{6}\phi^4\right],$$

$$P\doteq 0.60\phi\left[1+\frac{27}{120}\phi^4\right], \tag{19}$$

and Q approaches a value not equal to zero. It follows readily from this that $C_2=0$. Otherwise, v would not

[3] L. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1949), p. 181.

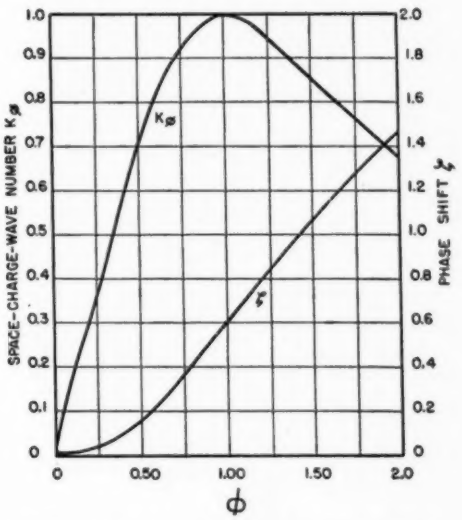


FIG. 3. Phase shift ζ and space-charge-wave number K_ϕ plotted against ϕ.

be finite at the cathode. Thus

$$J=C_1(P/v_0)\exp-j\omega\tau \tag{20}$$

and

$$v=\frac{C_1}{J_0}\left(\frac{-g^{\frac{1}{4}}}{j\omega}\frac{v_{02}}{3d}\right)\left(\frac{\partial\log v_0}{\partial\phi}-\frac{\partial\log P}{\partial\phi}\right)P\exp-j\omega\tau. \tag{21}$$

Thus from the initial condition for v, and using Eq. (19),

$$C_1=-\frac{1}{0.60}\left(\frac{j3\omega}{g^{\frac{1}{4}}}\frac{d}{v_{02}}\right)J_0v_1. \tag{22}\dagger$$

Now let J_{2p} and v_{2p} be the current and velocity modulations, respectively, at the anode that are obtained by the method neglecting geometry.[4] Thus

$$J_2=J_{2p}G_j,$$

$$v_2=v_{2p}G_v, \tag{23}$$

$$G_j=\frac{P_2}{0.60\phi_2},$$

$$\frac{G_v}{G_j}=\phi_2\left[\frac{\partial\log v_{02}}{\partial\phi}-\frac{\partial\log P_2}{\partial\phi}\right]. \tag{24}$$

Before undertaking any extensive calculations, it is useful to obtain values of G_j and G_v for small and large ϕ_2.

† This solution will satisfy the small-signal conditions everywhere except near the cathode. A rigorous solution can be obtained only by including the effect of thermal velocities.

[4] Cutler and Quate, Phys. Rev. **80**, 875 (1950).

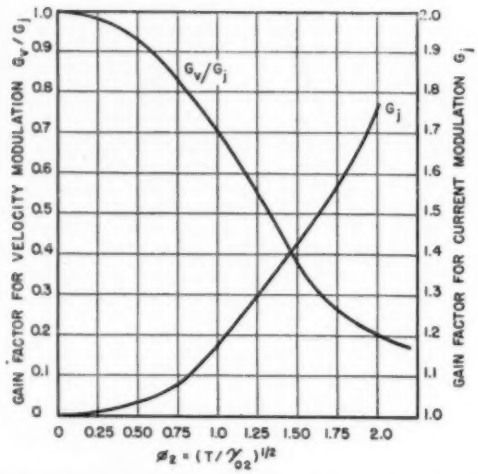

GAIN FACTOR FOR VELOCITY MODULATION G_v/G_j
GAIN FACTOR FOR CURRENT MODULATION G_j
G_v/G_j
G_j
0 0.25 0.50 0.75 1.0 1.25 1.50 1.75 2.0
$\phi_2=(T/\gamma_{02})^{1/2}$

FIG. 4. Gain factors for current and velocity modulations G_j and G_v/G_j plotted against $\phi_2=(T/\gamma_{02})^{\frac{1}{2}}$.

Small ϕ_2

Using Eq. (19), one obtains

$$\begin{aligned} G_j&=1+(27/120)\phi^4,\\ G_v/G_j&=1-(27/30)\phi^4. \end{aligned} \tag{25}$$

Large ϕ_2

For $\zeta_2>0.75$, one may use the asymptotic formulas for the Bessel function. Thus

$$P_2=\frac{1}{2.5}\frac{\exp\zeta_2}{K_2^{\frac{1}{2}}}, \tag{26}$$

$$G_j=\frac{P_2}{0.60\phi_2}, \tag{27}$$

$$\frac{G_v}{G_j}=2.5-\frac{\phi_2^4}{1+\phi_2^4}-\phi_2K_2. \tag{28}$$

Accurate values may be obtained by using Eq. (25) for $\phi_2<0.75$ and Eqs. (26)–(28) for $\phi_2\geq 1.25$. A curve of G_j and G_v/G_j *versus* ϕ_2 is plotted in Fig. 4. It is to be noted that the value of G_v/G_j for large ϕ_2 is independent of ζ_2.

(b) $G<1$ or $a>b$

Figure 4 may also be interpreted as a plot of G_j and G_v/G_j *versus* the value of $(T/\gamma_{02})^{\frac{1}{2}}$ at the anode. The simplest way to compute these quantities for this case would be to use Fig. 4 for the corresponding values at the anode. This would be rigorously true if G were a constant equal to $G(\phi_2)$. Physically this is plausible since the greatest effect of geometry would be at the anode where T/γ_0 attains its greatest value. G itself is a decreasing function of ϕ, which equals 1 for small values of ϕ and increases monotonically from 0 to 1 for large values of ϕ_2 as a/b goes from ∞ to 1. Now for large values of ζ_2, where the value of G_v/G_j is independent of ζ, only G (ϕ_2) is important. In this case and for small values of ϕ_2, this interpretation of Fig. 4 will be approximately correct. The curve of G_j will be somewhat in error for cases other than $G=1$. By the same reasoning, this result will also be approximately valid for a conical converging beam.

An experimental verification of this theory may be found in the experiments of Cutler and Quate[4] wherein the position of the current-modulation minimum in a drift tube following the diode did not agree with that computed from their theory. Using their notation,

$$\left[1+\left(\frac{T}{\gamma_{02}}\right)^2\right]^{\frac{1}{2}}=\frac{\omega_{p2}}{\omega(4QC^3)^{\frac{1}{2}}}. \tag{29}$$

According to their theory,

$$X_{\min}{}^{(1)} \text{ is proportional to } \tan^{-1}[\theta_1(4QC^3)^{\frac{1}{2}}]. \tag{30}$$

According to the theory presented here,

$$X_{\min}{}^{(2)}=\tan^{-1}\left[\theta_1(4QC^3)^{\frac{1}{2}}\frac{G_j}{G_v}\right]. \tag{31}$$

The numerical values in their experiment were $(4QC^3)^{\frac{1}{2}}=2.8\times10^{-2}$, $\omega_{p2}/\omega=0.052$, $T_2/\gamma_{02}=1.56$, $\phi_2=1.25$, $\theta_1=32$. From Fig. 4, $G_v/G_j=0.55$. Hence,

$$X_{\min}{}^{(2)}/X_{\min}{}^{(1)}=1.02/0.73=1.40. \tag{32}$$

By their measured values

$$X_{\min}{}^{(2)}/X_{\min}{}^{(1)}=1.1/0.7=1.57. \tag{33}$$

Thus the agreement is seen to be rather good. Furthermore, since G_v/G_j is always less than 1, the position of the minimum will always be farther from the anode than that from the theory neglecting geometry.

A better agreement in this case can be obtained by considering in greater detail the effect of the variation of G.

Thus in this case, Eq. (28) becomes, letting $\psi=G^{\frac{1}{4}}\phi$,

$$\frac{G_v}{G_j}=2.5+\frac{\phi_2}{4G}\frac{\partial G}{\partial\phi}-\frac{\psi_2^4}{1+\psi_2^4}\left[1+\frac{\phi_2}{4G}\frac{\partial G}{\partial\phi}\right]-\psi_2K(\psi_2). \tag{34}$$

Thus in the case $(\phi_2/4G)(\partial G/\partial\phi)\ll1$, the previous discussion of the effect of G is rigorously correct. For the case of $a/b=\infty$, which is so in these experiments, an approximate formula for G for large ϕ is

$$G=\frac{0.087}{0.09+\log\phi}. \tag{35}$$

In this case, $\phi_2=1.6$ from Fig. 1. Taking these values, $G_v/G_j=0.45$, which results in a computed value of $X_{\min}{}^{(2)}/X_{\min}{}^{(1)}$ equal to 1.52.

CONCLUSIONS

A general method has been given to obtain the effect of geometry on the propagation of space-charge waves in electron beams. The resulting differential equations, while complicated, may be approximately solved by WKB methods. The complete-space-charge diode has been discussed in detail, and the results have been checked by the experiments of Cutler and Quate. This method can be utilized for any dc conditions.

APPENDIX. FIELD-THEORY ANALYSIS OF PROPAGATION OF SPACE-CHARGE WAVES IN FINITE BEAMS

An attempt will now be made to derive the method of evaluating the effect of geometry from field-theory considerations. In this case, the ballistic equation is

$$v_0^3\frac{\partial^2 J}{\partial Z^2}+\left[3v_0^2\frac{\partial v_0}{\partial Z}+2j\omega v_0^2\right]\frac{\partial J}{\partial Z}+\left[2j\omega v_0\frac{\partial v_0}{\partial Z}-\omega^2 v_0\right]J=\frac{eJ_0}{m}j\omega E. \quad (36)$$

Here E is the component of the electric field in the direction of propagation. All transverse motions are neglected. The derivation of this is similar to that of Eq. (7).

The circuit equation, which formerly was given by Eq. (5), is now

$$\nabla_t^2 E+\frac{\partial^2 E}{\partial Z^2}+k^2E=j\frac{\zeta}{k}\left(k^2J+\frac{\partial^2 J}{\partial Z^2}\right), \quad (37)$$

where

$$\nabla_t^2=\frac{\partial^2}{\partial X^2}+\frac{\partial^2}{\partial Y^2},$$

$k=2\pi/\lambda$, the free-space wave number, and $\zeta=(\mu_0/\epsilon_0)^{\frac{1}{2}}$.

This is derived directly from the wave equation. The case of $a=b$ will be considered first. Here let

$$E=E_1\exp(-j\omega\tau)J_0(Tr),$$
$$J=J_1\exp(-j\omega\tau)J_0(Tr),$$
$$\text{Ta}=2.4. \quad (38)$$

This solution will satisfy the boundary condition that $E=0$ for $r=a$. E_1 is a function of Z only.

Equations (36) and (37) now become

$$v_0\frac{\partial^2 J_1}{\partial\tau^2}+2\frac{\partial v_0}{\partial\tau}\frac{\partial J_1}{\partial\tau}=\frac{e}{m}-J_0j\omega E_1,$$

$$v_0\frac{\partial^2 E_1}{\partial\tau^2}-\left(\frac{\partial v_0}{\partial\tau}+2j\omega v_0\right)\frac{\partial E_1}{\partial\tau}+\left[(-\gamma_0^2-T^2+k^2)v_0^3+j\omega\frac{\partial v_0}{\partial\tau}\right]E_1=j\frac{\zeta}{k}\left[v_0\frac{\partial^2 J_1}{\partial\tau^2}-\left(\frac{\partial v_0}{\partial\tau}+2j\omega v_0\right)\frac{\partial J_1}{\partial\tau}+\left[(-\gamma_0^2+k^2)v_0^3+j\omega\frac{\partial v_0}{\partial\tau}\right]J_1\right]. \quad (39)$$

Now it is known from the infinite-beam solution that E_1 varies as $\exp\pm j\omega_p\tau$. Hence if one assumes that:

$$\omega_p/\omega\ll 1, \quad (41)$$

$$\frac{1}{\omega v_0}\frac{\partial v_0}{\partial\tau}\ll 1, \quad (42)$$

$$k/\gamma_0\ll 1, \quad (43)$$

then the terms containing the derivatives may be neglected in Eq. (40). Equation (40) reduces to

$$E_1=-j\frac{\zeta}{k}\frac{J_1}{1+(T^2/\gamma_0^2)}. \quad (44)$$

Substituting this in Eq. (39), one has

$$v_0\frac{\partial^2 J_1}{\partial\tau^2}+2\frac{\partial v_0}{\partial\tau}\frac{\partial J_1}{\partial\tau}+\frac{e}{m\epsilon_0}\frac{J_0}{1+(T^2/\gamma_0^2)}J_1=0. \quad (45)$$

Thus it is seen that J_0 is reduced by the factor as stated in Eq. (2). This approximation will also be valid in the case $a>b$ provided T is a slowly varying function of Z. Conditions for Eqs. (41) and (42) are obviously not satisfied near the cathode of a complete-space-charge diode. However, this does not matter as the effect of geometry is small there.

The Dielectric Constant of Water Vapor in the Microwave Region*

GEORGE BIRNBAUM
National Bureau of Standards, Washington, D. C.

AND

S. K. CHATTERJEE†
Indian Institute of Science, Bangalore, India

(Received August 20, 1951)

The dielectric constant of water vapor was determined at 9280 Mc by a cavity comparison method at 10 temperatures ranging from 32 to 103°C. In addition, observations were made at 24,800 Mc at the single temperature of 24.5°C. Although the vapor pressure was raised to within 10 percent of the saturation value at 24.5°C and 103°C, the variation of $(\epsilon'-1)/(\epsilon'+2)$ with pressure remained linear, contrary to observations made at radiofrequencies. These data are discussed in relation to the questions of association and adsorption of the vapor. The present results are in agreement with those obtained at radiofrequencies (sufficiently below saturation pressures) and do not show the dispersion reported in an early microwave experiment. Results obtained here and those obtained by others are used to determine mean values for the constants in the Debye equation: namely, $P=[3.96\pm0.32]+[(2.077\pm0.016)10^4/T]$. This gives a dipole moment of $(1.846\pm0.005)10^{-18}$ esu.

1. INTRODUCTION

FROM the well-known theory of Debye, the molar polarization P is given by

$$P=(\epsilon'-1)RT/(\epsilon'+2)p_i=A+B/T.$$

Deviations from this equation near the saturation point of water vapor have been observed in measurements of dielectric constant at radio frequencies.[1] Such deviations have been attributed to association of vapor molecules and effects attendant on adsorption of vapor on the condenser system.[2] Early in the development of microwave techniques, Saxton[3] determined the dielectric constant of steam above 100°C and at atmospheric pressure, and found a dispersion in the frequency region 3000 to 18,000 Mc which he was unable to account for theoretically.

On account of the previously mentioned anomalies, we decided to measure the dielectric constant of water vapor over a rather wide range of pressures and temperatures by an accurate microwave method. The results obtained from our experiments and those obtained from a critical survey of the literature were used to arrive at mean values for the constants A and B in the Debye equation. These results should be useful, for example, in connection with accurately determining the velocity of propagation of radio waves. They will be used here to obtain the dipole moment and to reveal a discrepancy between the experimentally determined polarization extrapolated to $T=\infty$ and the sum of the optical and infrared polarizations.

2. EXPERIMENTAL ARRANGEMENT

The experiments were performed at 9280 and 24,800 Mc with a cavity comparison method fully described in a previous paper;[4] therefore, a description of only the new experimental features will suffice here. Standard transmission-type cavities, which oscillate in the circular TE_{011} mode with a loaded Q of 1×10^4, were employed at 9280 Mc. A ring of polyiron for damping oscillations that might be present in the region behind the tuning plunger was hygroscopic and accordingly removed. Tests showed that the normal behavior of the cavity, at least at the present frequency, was unaffected by this change. Mechanical distortion of the test cavity with evacuation was prevented by enclosing it in an appropriate chamber.[5] The temperature of the test cavity was determined to within 0.1°C by an arrangement consisting of a mercury thermometer embedded in the wall of the evacuation chamber and thermocouples connected between the chamber and cavity within. At 24,800 Mc, cylindrical TE_{013} cavities based on the design of Bleaney *et al.*[6] were used. Tests of these cavities with a gas of known $\epsilon'-1$ showed that no evacuation chamber was necessary for work in the low pressure region of interest.

To realize the convenience of measuring the pressure of the hot vapor with an accurate mercurial barometer kept at room temperature, we employed an apparatus described by Zahn and Miles.[7] After some vapor pres-

* A preliminary account of some of the work reported here was presented at the 1950 Conference on Electrical Insulation, Pocono Manor, Pennsylvania. An abstract was published in the 1950 Annual Report of the Conference on Electrical Insulation, National Research Council.

† Formerly a guest worker at the National Bureau of Standards.

[1] A review of reports of such deviations is given by J. D. Stranathan, Phys. Rev. **48**, 538 (1935). See also L. H. Ford, J. Inst. Elec. Engrs. (London) **95**, Part II, 709 (1948).

[2] For further discussion of these matters see C. T. Zahn, Phys. Rev. **35**, 1047 (1930).

[3] J. A. Saxton, "The dielectric properties of water vapor at very high frequencies," in *Meteorological Factors in Radio-Wave Propagation* (published by the Physical Society, London, 1946), p. 215.

[4] Birnbaum, Kryder, and Lyons, J. Appl. Phys. **22**, 95 (1951).

[5] G. Birnbaum, Rev. Sci. Instr. **21**, 173 (1950). See Fig. 6.

[6] Bleaney, Loubser, and Penrose, Proc. Phys. Soc. (London) **59**, 185 (1947).

[7] J. B. Miles, Jr., Phys. Rev. **34**, 964 (1929); C. T. Zahn, Rev. Sci. Instr. **1**, 299 (1930). If accurate low pressure measurements are to be made, it should be noted that at the equilibrium point of the apparatus the level of liquid is higher in the narrower part of the U-tube; this, of course, is due to capillarity.

sure had been apparently established it was noticed, especially at lower pressures, that the pressure of vapor in the cavity decreased and the cavity frequency increased. The same changes were also observed with a simpler vapor system, used for some preliminary experiments, after the vapor supply was cut off. These effects, presumably due to adsorption, occurred although vapor was admitted into the cavity quite slowly.

The probable error of a single $\epsilon'-1$ determination is about 0.6 percent or 1×10^{-6}, whichever is greater. This estimate includes the statistical errors in measuring pressure and temperature as well as the statistical error in measuring changes in frequency. A systematic error of roughly $\pm0.7\times10^{-6}$ was possible, due chiefly to the measurement of pressure.

3. RESULTS AND DISCUSSION

At 9280 Mc the variation of $\epsilon'-1$ with pressure was determined at temperatures ranging from 32.0°C to 103.4°C, while at 24,800 Mc the data were obtained only at 24.5°C. For the sake of clarity only part of the data is shown in Fig. 1. The pressure the vapor would exert were it an ideal gas is represented by p_i; it is obtained from the observed pressure p and the van der Waals constants a and b, thus

$$p_i=p\{1+(p/RT)[(a/RT)-b]\}.$$

Except at pressures near 1 atmos (100°C) where the deviation from ideal behavior was about 0.5 percent, the deviations at all other pressures attained in this experiment were less than 0.1 percent.

The important feature of the curves in Fig. 1 is the absence of nonlinearities. On the other hand, various observations of an anomalous pressure dependence of $(\epsilon'-1)/(\epsilon'+2)$ have been reported in measurements at radio frequencies. Stranathan[1] investigated this matter and concluded that the deviations were due wholly to the added polarization contributed by the film of water adsorbed on the insulating supports of the condenser. He was able to reduce this effect by placing the insulators in the weak portion of the electric field; nevertheless a deviation from linearity set in as the saturation pressure was approached. In view of this, it is worth while to point out that no insulating supports are present in cavities. In addition, although vapor was undoubtedly adsorbed by the walls of the cavity in the present experiment,[8] the perturbation of the resonance frequency is negligible because the electric field is essentially zero on the cavity walls for the mode of oscillation used.

It is easy to see that if association exists, it would in general give rise to a nonlinear variation of molecular polarization with pressure. Such a variation would also be observed, however, if dielectric measurements were

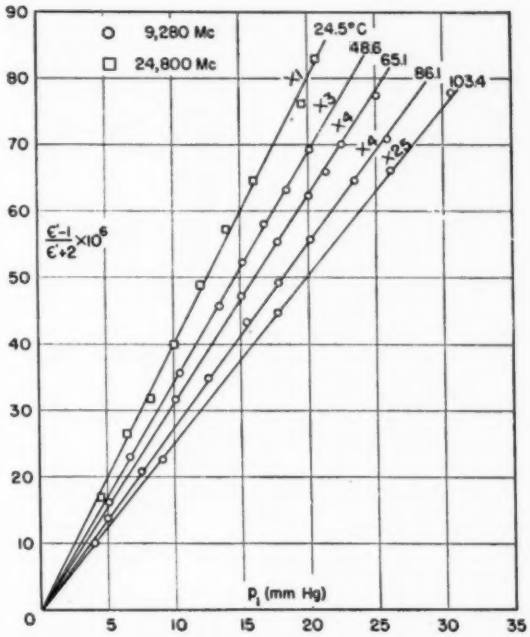


FIG. 1. The quantity $(\epsilon'-1)/(\epsilon'+2)$ plotted against pressure p_i in mm Hg. Both ordinate and abscissa should be multiplied by the factor indicated.

made in a region of strong spectral absorption.[9] This case is easy to identify because significant absorption of radiation by the vapor would cause an easily observed lowering of the cavity Q. With this in view, it is apparent from the considerations of the previous paragraph that the cavity method is well suited for the study of association.

The linearity of the curves in Fig. 1, especially those for 24.5 and 103.4°C where the pressure was raised to within 10 percent of saturation, indicates the absence of any appreciable association. This conclusion is in agreement with that derived from vapor density measurements. Indeed, if there is association even at saturation the amount cannot be great, for the total departure of the measured density from that of an ideal gas is only about 1.5 percent at 760 mm Hg, 100°C.[10]

Using all the experimental points, at each temperature the polarization was determined from the slope of the appropriate curve in Fig. 1. This polarization plotted against $1/T$(°K) is shown in Fig. 2. The straight line was determined from a least squares treatment of the 9280 Mc data, and it is noteworthy that the value of polarization obtained at 24,800 Mc falls exactly on this line.

A least squares solution for the best straight line for all the points in Fig. 2 gives

$$(\epsilon'-1)RT/(\epsilon'+2)p_i=[3.84\pm0.72]+[(2.092\pm0.024)10^4/T],$$

[8] The effect of adsorption of water vapor on the walls of a very large electromagnetic cavity is given by G. E. Becker and S. H. Autler, Phys. Rev. **70**, 305 (1946).

[9] An example of the effect of strong absorption lines on $(\epsilon'-1)-p$ data is given by G. Birnbaum, Phys. Rev. **77**, 144 (1950).

[10] This can be seen by examining the data of Osborne, Stimson, and Ginnings, NBS J. Research **23**, 261 (1939).

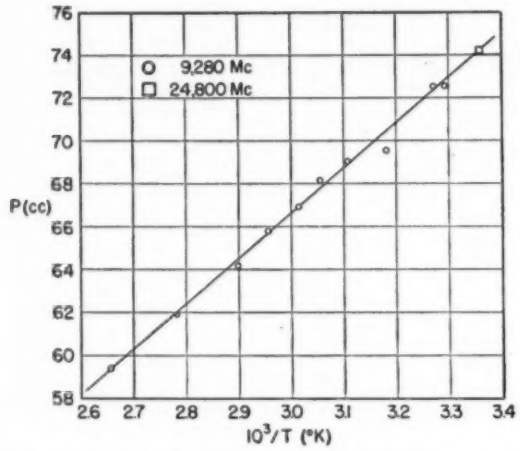


FIG. 2. The polarization plotted against $1/T(°K)$. The straight line was determined from a least squares treatment of the 9280 Mc data.

where the uncertainties are probable errors. The values of polarization were weighted, lower weights being assigned to the values at the lower temperatures. In this temperature region it was necessary to use lower pressures which gave a larger relative error in the pressure and frequency measurements. For comparison we give the constants obtained by using the unweighted data, *viz.*, $A=4.3\pm1.1$ and $B=(2.076\pm0.035)10^4$. The error in T was taken to be negligible in comparison with the error in P in the least squares treatment of the $P-T$ data discussed here and elsewhere in this paper.

The present results are in satisfactory agreement with those of Stranathan[1] at 1 Mc, Phillips[11] at 3000 Mc, Crain[12] at 9000 Mc, and Essen and Froome[13] at 24,000 Mc. Phillips and Essen made measurements at room temperature only. Crain's data were taken from 20 to 68°C over a limited range of pressure. There is a weak rotational absorption line at 1.35 cm, but Van Vleck[14] has shown that this is not expected to alter the dielectric constant appreciably. One of us (G.B.) has completed a preliminary calculation of dispersion, based on the effect of the entire rotational spectrum of water vapor. This shows, for example, that the dipole term in the Debye equation (at room temperature) increases only by about 1 percent at a wavelength of 3 mm.

4. MEAN VALUES OF A AND B IN THE DEBYE EQUATION

Although many observations have been reported on the dielectric constant of water vapor, few have been taken over a sufficient range of temperatures and pressures to allow an accurate determination of A and B. Those obtained from a critical survey of the literature and also our values are shown in Table I.[15] We calculated the values assigned to Groves and Sugden[16] and Hurdis and Smyth[17] from their data. The weighted means of A and B obtained from the values listed in Table I are

$$A=3.96\pm0.32,$$

$$B=(2.077\pm0.016)10^4.$$

As is customary, the weights were taken inversely proportional to the square of the probable errors.

Thus far we have ignored the fact that the Debye equation is an asymptotic formula valid at high temperatures. This has been shown from quantum-mechanical considerations by Van Vleck[18] who has approximately corrected the Debye formula by adding a term of the form $-C/T^2$, where at room temperature $C/T\ll B$. In the interval of temperature under consideration, however, the variation in this term is small enough so that the experimental data are accurately represented by the Debye equation.

Using the values of B given in Table I, the mean value of dipole moment was found to be 1.846 ± 0.005 D. The polarization at $T=\infty$ has been given above as 3.96 ± 0.32. A rather lower value for this quantity and a somewhat higher value for the dipole moment would be obtained if the data were analyzed in the light of the modified Debye equation. This matter is briefly discussed in the Appendix.

We wish to thank Mr. A. Hanyok for assistance in taking the data. We also wish to express our appreciation to Dr. A. A. Maryott for several valuable discussions and to Dr. H. Lyons for his interest in this work.

APPENDIX

The modified Debye equation is

$$P=A'+[4\pi N_0\mu^2/9kT][1-(K/T)], \quad (1)$$

where K (which depends in part on the moments of inertia of the molecule) has the value 8.8 for the water molecule, and all the

TABLE I. Values of the constants A and B in the Debye equation. The errors are probable errors.

Observer	A	$B\times10^{-4}$	Temperature range °C
Birnbaum and Chatterjee	3.84 ± 0.72	2.092 ± 0.024	25–103
Groves and Sugden[a]	4.3 ± 1.2	2.074 ± 0.054	110–211
Hurdis and Smyth[b]	3.4 ± 1.2	2.087 ± 0.053	111–249
Stranathan[c]	4.03 ± 0.39	2.071 ± 0.014	21–189

[a] See reference 16.
[b] See reference 17.
[c] See reference 1.

[11] W. C. Phillips, Proc. Inst. Radio Engrs. **38**, 786 (1950).
[12] C. M. Crain, Phys. Rev. **74**, 691 (1948).
[13] L. Essen and K. D. Froome, Nature **167**, 512 (1951).
[14] J. H. Van Vleck, "The relation between absorption and dispersion," in *Propagation of Short Radio Waves* edited by D. E. Kerr, M.I.T. Radiation Laboratory Series Vol. 13 (McGraw-Hill Book Company, Inc., New York, 1951), p. 641.

[15] The work of R. Sanger and O. Steiger, Helv. Phys. Acta **1**, 369 (1928), and R. Sanger, Physik. Z. **31**, 306 (1930) is not listed here. Their values ($A=3.9$, $B=2.09\times10^4$), which are in close agreement with those in Table I, to be sure, are based on only 3 points.
[16] I. G. Groves and S. Sugden, J. Chem. Soc. (London), 971 (1935).
[17] E. C. Hurdis and C. P. Smyth, J. Am. Chem. Soc. **64**, 2829 (1942).
[18] J. H. Van Vleck, *The Theory of Electric and Magnetic Susceptibilities* (Oxford University Press, London, 1932), p. 198.

other symbols have the usual meanings. The factor $[1-(K/T)]$ changes the values of electric moment and P at $T=\infty$ as obtained by fitting the experimental data to the Debye equation. Making the changes in the manner described by Van Vleck[18] and taking $T=375°$A as a mean value, we find that $\mu=1.889D$ and $A'=2.76\pm0.32$. At $T=\infty$, P should equal the sum of the optical and infrared polarizations extrapolated to zero frequency. The former was found to be 3.59 ± 0.03 from the work of Barrell and Sears[19] and the latter 0.35 from the work of Greenfield and Brown.[20] The sum in question is then 3.94 and differs from $A'=2.76$ by an amount which appears to be outside experimental errors. The cause of this discrepancy is not known.

[19] H. Barrell and J. E. Sears, Trans. Roy. Soc. (London) **238A**, 59 (1939–1940).

[20] M. A. Greenfield and F. W. Brown, J. Opt. Soc. Am. **40**, 643 (1950).

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 2 FEBRUARY, 1952

Generalized Ideal Filters

LOTFI A. ZADEH AND KENNETH S. MILLER
Columbia University and New York University, New York City

(Received August 8, 1951)

One of the basic concepts in the theory of electrical filters is that of an ideal filter. As commonly defined, an ideal filter is a network which passes without distortion all frequency components falling within a certain frequency range while rejecting all other frequency components. In the present paper a broader definition is formulated which extends the concept of ideal filter to both linear varying-parameter and nonlinear types of systems. Briefly, a filter N is said to be ideal if it can extract a set of signals $\mathfrak{M}$ from the sum of $\mathfrak{M}$ and some other set $\mathfrak{N}$. A direct consequence of this definition is that any ideal filter N (linear or nonlinear) is idempotent, that is, is equivalent to a tandem combination of two filters each of which is identical with N. The converse, however, is true only in the case of linear filters.

The basic properties of ideal filters are investigated by the use of function space techniques. It is shown, in particular, that by employing linear ideal filters one can separate two (or more) simultaneously transmitted sets of signals which occupy overlapping frequency bands, provided only the sets in question span disjoint manifolds in the signal space.

I. INTRODUCTION

IN the theory of electrical networks an ideal filter is commonly understood to be a system which, loosely speaking, passes without distortion all frequency components falling within a certain range (the pass-band of the filter) and rejects all frequency components falling outside this range. Such filters, although unrealizable in an exact manner in the form of lumped-constant networks, play an important role in the theory of filter design and have been the subject of numerous investigations.[1,2]

In the present paper a broader definition of ideal filter which applies to both linear varying-parameter and nonlinear types of systems is formulated; and a theory of linear ideal filters, based largely on the use of function space methods and other techniques commonly used in quantum mechanics,[3] is outlined.

Preliminary investigation indicates that the use of "generalized" ideal filters affords some novel and advantageous modes of approach to the problems of simultaneous transmission and filtration of signals. In particular, it will be seen that by using linear ideal filters it is possible to separate two or more simultaneously transmitted sets of signals which occupy overlapping frequency bands, provided only that the sets in question span disjoint manifolds in the signal space.

[1] H. W. Bode, J. Math. and Phys. **13**, 275 (1934).

[2] E. A. Guillemin, *Communication Networks* (John Wiley and Sons, Inc., New York, 1935), Vol. II.

[3] H. Weyl, *Theory of Groups and Quantum Mechanics* (Dover Publications, Inc., New York, 1949).

2. FUNDAMENTALS

It will be convenient to identify a filter (not necessarily ideal or linear) by the symbol N, and to write

$$v=Nu \tag{1}$$

in order to signify that the signals $u(t)$ and $v(t)$ constitute, respectively, the input and output of N. Except where otherwise specified, it will be assumed that N is initially at rest.

In the analysis that follows, it will be helpful to use the function space representation of signals. Thus, a signal $u(t)$ will be represented as a vector u in a function space (signal space) Σ. A Hilbert space, possibly of nonseparable type, is adequate for most practical purposes. For random signals, it is necessary to introduce a probability measure into the signal space. A theory of random processes based on Hilbert space representation has been given by Karhunen.[4]

Consider two (or more) sets of signals, $\{\mu(t)\}=\mathfrak{M}$ and $\{\nu(t)\}=\mathfrak{N}$. $\mathfrak{M}$ and $\mathfrak{N}$ are two specified classes of signals and form (not necessarily linear) manifolds in the signal space Σ. It will be assumed that both $\mathfrak{M}$ and $\mathfrak{N}$ include the null element of the space.

[4] K. Karhunen, Ann. Acad. Fennicae **34**, AI (1946).

A generalized ideal filter, or simply an ideal filter, will be defined in the following manner:

Definition

A filter N (linear or nonlinear) will be said to be *ideal* if for any $\mu(t)$ in $\mathfrak{M}$ and any $\nu(t)$ in $\mathfrak{N}$,

$$N(\mu+\nu)=\mu. \tag{2}$$

Since both $\mathfrak{M}$ and $\mathfrak{N}$ include the null element, Eq. (2) implies that

$$N\mu=\mu \quad \text{for} \quad \mu \text{ in } \mathfrak{M}, \tag{3}$$

$$N\nu=0 \quad \text{for} \quad \nu \text{ in } \mathfrak{N}. \tag{4}$$

This means that N passes without distortion all signals belonging to $\mathfrak{M}$ and rejects all those belonging to $\mathfrak{N}$. On this basis, the manifolds $\mathfrak{M}$ and $\mathfrak{N}$ will be referred to respectively as the *acceptance* and *rejection* manifolds of N.

It is to be noted that Eqs. (3) and (4) do not necessarily imply Eq. (2), except in the case of linear filters. Thus, a nonlinear filter may satisfy Eqs. (3) and (4) for some $\mathfrak{M}$ and $\mathfrak{N}$ without being an ideal filter. For example, a linear rectifier satisfies Eqs. (3) and (4)where $\mathfrak{M}$ is the set of all non-negative signals and $\mathfrak{N}$ is the set of all nonpositive signals. Yet Eq. (2) clearly does not hold for any positive $\mu(t)$ and negative $\nu(t)$. Consequently, a linear rectifier is not an ideal filter.

A fundamental characteristic of ideal filters is their *idempotency*. More specifically, if N is an ideal filter, then a tandem combination of two such filters is equivalent to N. (The equivalence is supposed to hold for a specified class of input signals only.) This property of ideal filters is a direct consequence of the defining Eq. (2). Thus, consider a tandem combination of two ideal filters each of which is identical with N. From Eq. (2), the response of the first member of this combination to $\mu+\nu$ is μ. The response of the second member to μ is μ. Consequently, the response of the tandem combination to $\mu+\nu$ is μ—which demonstrates that N followed by N is equivalent to N.

An important question which arises in connection with the idempotent property of ideal filters is that of the sufficiency of this property. In other words, is a filter ideal if it is idempotent? The answer to this question is "yes" in the case of linear filters and "no" in the case of nonlinear filters. This is seen by the following reasoning.

First, suppose that N is an idempotent linear filter. Let $v(t)$ be the response of N to $u(t)$, that is, $Nu=v$. Since by hypothesis $N(Nu)=Nu$, it follows that $Nv=v$. This implies that there is a class of signals $\{v\}$ which is passed by N without distortion. (The class $\{v\}$ forms a linear manifold $\mathfrak{M}$ in Σ.) Also, from $Nu=v$ follows $N(v-u)=0$. This shows that there is a class of signals, $\mathfrak{N}=\{v-u\}$, which is rejected by N. Consequently, in view of Eqs. (3) and (4) and the linearity of N, N is an ideal filter. Thus, *any idempotent linear filter is an ideal filter*. This conclusion can be arrived at indirectly (as will be seen later) by using the analogy between linear ideal filters and projections in linear vector spaces.

In the case of nonlinear filters, an idempotent filter is not necessarily an ideal filter. A linear rectifier is obviously an idempotent filter, and yet, as was shown earlier, it is not an ideal filter.

Before proceeding further with the investigation of various properties of ideal filters, it will be convenient to examine a few examples in the light of the foregoing discussion.

First, consider a conventional ideal filter F, say of the low pass type, which passes all frequencies below f_0 cps and rejects all those above f_0 cps. The gain of this filter is equal to unity over the pass-band $|f|\leqq f_0$ and is equal to zero for $|f|>f_0$. The phase is zero throughout. The acceptance manifold $\mathfrak{M}$ consists of the class of signals not containing frequencies above f_0 cps while the rejection manifold consists of the class of signals not containing frequencies below or equal to f_0 cps. It is clear that F is idempotent and that it is an ideal filter in the sense of Eq. (2).

Next consider a two-pole switch S which is closed over the interval $0\leqq t\leqq T$ and is open for $t<0$ and $t>T$. It is clear that S is a linear idempotent system, and consequently it is an ideal filter. The acceptance manifold of S consists of the class of signals which vanish outside the interval $0\leqq t\leqq T$, and the rejection manifold is made up of signals which vanish over this interval.

Finally, consider the decoder of a PCM (pulse code modulation) system. The decoder may be regarded as an ideal nonlinear filter whose acceptance manifold consists of trains of binary (on-off) pulses of specified width and amplitude, and whose rejection manifold is the class of signals whose amplitude is less than half of the quiescent pulse amplitude. It is clear that the decoder is an idempotent system.

In the following sections of this paper we shall confine our attention to linear filters inasmuch as the analytical machinery for dealing with nonlinear filters has not yet been sufficiently well developed.[5]

3. LINEAR FILTERS

In the case of linear ideal filters the acceptance and rejection manifolds $\mathfrak{M}$ and $\mathfrak{N}$ have the following properties.

(i) $\mathfrak{M}$ and $\mathfrak{N}$ are linear [6] manifolds.
(ii) $\mathfrak{M}$ and $\mathfrak{N}$ are disjoint. That is, the only signal common to both $\mathfrak{M}$ and $\mathfrak{N}$ is the null signal.
(iii) $\mathfrak{M}$ and $\mathfrak{N}$ are complementary. That is, if u is any signal in Σ then it can be resolved in one and only one way into two component signals μ and ν lying in $\mathfrak{M}$ and $\mathfrak{N}$ respectively.

[5] Further discussion of nonlinear filters is given in reference 11.
[6] F. J. Murray, *Linear Transformations in Hilbert Space* (Princeton University Press, Princeton, 1941).

Note that (iii) implies (ii). In function space notation we write for linear manifolds,

$$\mathfrak{M}\cdot\mathfrak{N}=0 \tag{5}$$

$$\mathfrak{M}\oplus\mathfrak{N}=\Sigma, \tag{6}$$

that is, the intersection of $\mathfrak{M}$ and $\mathfrak{N}$ is the zero element, and the direct sum of $\mathfrak{M}$ and $\mathfrak{N}$ is the whole space. Equations (5) and (6) together imply that $\mathfrak{M}$ and $\mathfrak{N}$ are complementary manifolds.

Consider a signal u in Σ and let μ and ν be the components of u in $\mathfrak{M}$ and $\mathfrak{N}$, respectively. In function space terminology, the operator E defined by the relation

$$E(\mu+\nu)=\mu \tag{7}$$

defines the projection of Σ on $\mathfrak{M}$ along $\mathfrak{N}$. From a comparison of Eqs. (7) and (2) it is clear that, operationally, E is equivalent to N. This means, in physical terms, that the projection of Σ on $\mathfrak{M}$ along $\mathfrak{N}$ is equivalent to a process of filtration using an ideal filter N whose acceptance and rejection manifolds constitute $\mathfrak{M}$ and $\mathfrak{N}$, respectively.

The analogy between projection in a function space and filtration with an ideal filter is useful in two respects. First, in dealing with ideal filters it is frequently helpful to employ a geometrical point of view based on this analogy and the geometrical properties of projections. Second, the analogy makes it possible to apply to the study of ideal filters the vast amount of available information about the various properties of projections in a function space. For example, it is well, known that a projection is an idempotent operator and conversely, an idempotent operator (defined on Hilbert space) is a projection. Translated into terms of ideal filters this means that any linear idempotent filter is an ideal filter—a conclusion which was arrived at earlier by a direct mode of reasoning.

Many other properties of ideal filters follow from the well-known properties of algebraic combinations of projections. Preliminary to a discussion of some of these properties it will be expedient to adopt a form of functional symbolism which is illustrated in Fig. 1. Thus, $N=N_1+N_2$ represents a filter N consisting of two filters N_1 and N_2 connected in the manner shown in Fig. 1(a). (The circle with a plus sign inside represents an "adder.") Similarly, $N=N_2N_1$ (Fig. 1(b)) represents a tandem combination of N_2 and N_1, with N_1 operating on the input first and N_2 operating on the output of N_1. The symbol 1 denotes a *unit* filter, that is, a filter whose output is identical with the input for all inputs. The symbol N^{-1} represents the inverse of N (when it exists), that is, N^{-1} is a filter such that a tandem combination of N and N^{-1}, or of N^{-1} and N, is equivalent to a unit filter. The notation $N=N_1|N_2$, which will be used at a later point, has the significance illustrated in Fig. 1(c).

The salient properties of several modes of combination of ideal filters, which follow from the corresponding

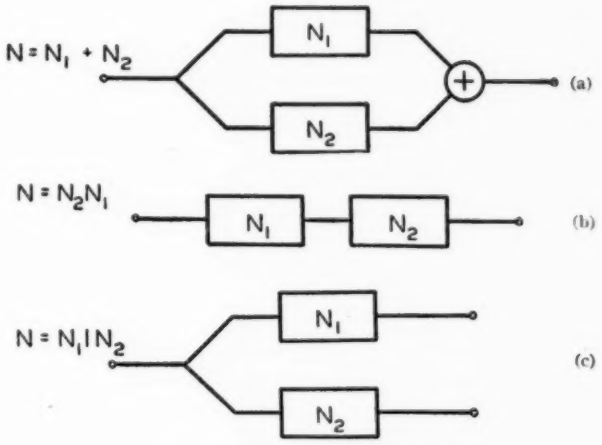


FIG. 1. Functional symbolism.

properties of projections, are listed below. In the following statements the symbols N_1 and N_2 denote two ideal filters whose acceptance and rejection manifolds are designated by $\mathfrak{M}_1$, $\mathfrak{N}_1$ and $\mathfrak{M}_2$, $\mathfrak{N}_2$, respectively. $\mathfrak{M}$ and $\mathfrak{N}$ denote the acceptance and rejection manifolds of the combination of N_1 and N_2. The symbols $\cdot$ and $\oplus$ are used to indicate the intersection and direct sum of two manifolds. The proofs of these statements (in terms of projections) may be found in any text on linear vector spaces.[6]

(1) If N is an ideal filter whose acceptance and rejection manifolds are $\mathfrak{M}$ and $\mathfrak{N}$, respectively, then $1-N$ is an ideal filter whose acceptance manifold is $\mathfrak{N}$ and whose rejection manifold is $\mathfrak{M}$. The ideal filter $1-N$ is said to be *complementary* to N.

(2) The filter $N=N_1+N_2$ is an ideal filter if and only if $N_1N_2=N_2N_1=0$. If this condition is satisfied, the acceptance and rejection manifolds of N are given by $\mathfrak{M}=\mathfrak{M}_1\oplus\mathfrak{M}_2$ and $\mathfrak{N}=\mathfrak{N}_1\cdot\mathfrak{N}_2$.

(3) The filter $N=N_1-N_2$ is an ideal filter if and only if $N_1N_2=N_2N_1=N_2$. For this case $\mathfrak{M}=\mathfrak{M}_1\cdot\mathfrak{N}_2$ and $\mathfrak{N}=\mathfrak{N}_1\oplus\mathfrak{M}_2$.

(4) The filter $N=N_1N_2$ is an ideal filter if $N_1N_2=N_2N_1$. In this case $\mathfrak{M}=\mathfrak{M}_1\cdot\mathfrak{M}_2$ and $\mathfrak{N}=\mathfrak{N}_1\oplus\mathfrak{N}_2$.

An important property of ideal filters relates to their invariance under linear transformations. Thus:

(5) If N is an ideal filter and L is a nonsingular filter (that is, one that has an inverse), then LNL^{-1} is an ideal filter. The acceptance and rejection manifolds of this filter are, respectively, $L(\mathfrak{M})$ and $L(\mathfrak{N})$, where $\mathfrak{M}$ and $\mathfrak{N}$ are the corresponding manifolds of N, and $L(\mathfrak{M})$ and $L(\mathfrak{N})$ represent the manifolds resulting from operating with L on the elements of $\mathfrak{M}$ and $\mathfrak{N}$, respectively.

This statement can readily be proved by showing first that LNL^{-1} is an idempotent filter. We have

$$(LNL^{-1})(LNL^{-1})=LN^2L^{-1}=LNL^{-1}, \tag{8}$$

which demonstrates that LNL^{-1} is idempotent and

hence an ideal filter. Next from

$$\begin{matrix} N(\mathfrak{M})=\mathfrak{M} \\ N(\mathfrak{N})=0 \end{matrix} \tag{9}$$

we conclude

$$LNL^{-1}[L(\mathfrak{M})]=L(\mathfrak{M}) \tag{11}$$

$$LNL^{-1}[L(\mathfrak{N})]=L[N(\mathfrak{N})]=0, \tag{12}$$

which establishes the second portion of the statement. It should be remarked that the statement in question (5) holds also when N is a nonlinear ideal filter.

4. ANALYTIC FORMULATION[7]

In working with ideal filters at a practical level it is necessary to deal with analytical formulations of the various relations pertaining to linear filters and, more particularly, ideal linear filters. Such formulations can readily be obtained by introducing a suitable basis (coordinate system) into the signal space Σ. This is equivalent to expressing the signal $u(t)$ in Σ in an integral form

$$u(t)=\int_C k(t,\lambda)U(\lambda)d\lambda, \tag{13}$$

where λ is, in general, a complex parameter; $k(t,\lambda)$ is a suitable set of "component signals" in terms of which $u(t)$ is resolved; $U(\lambda)$ is the spectral function of $u(t)$ relative to $k(t,\lambda)$; and C is a contour in the λ-plane (generally a straight line parallel to either the imaginary or real axis). [When the $k(t,\lambda)$ form a discrete set (for example, complex Fourier series), the spectral function $U(\lambda)$ assumes the form of a train of delta-functions. This is a more convenient viewpoint than to write Eq. (13) as a Stieltjes integral.] It is assumed that the kernel $k(t,\lambda)$ has an inverse $k^{-1}(\lambda,t)$ such that

$$\int_C k(t,\lambda)k^{-1}(\lambda,\xi)d\lambda=\delta(t-\xi), \tag{14}$$

where $\delta(t-\xi)$ is a delta-function. In terms of $k^{-1}(\lambda,t)$ the spectral function $U(\lambda)$ may formally be written as

$$U(\lambda)=\int_{-\infty}^{+\infty} k^{-1}(\lambda,t)u(t)dt. \tag{5}1$$

The set of functions $k(t,\lambda)$ constitute a basis for Σ (or for a space in which Σ is imbedded). For convenient reference, a few basic types of $k(t,\lambda)$ kernels and their respective inverses are listed below. For brevity, the various well-known types of discrete orthogonal sets are omitted.

(1) $k(t,\lambda)=\delta(t-\lambda)$; $k^{-1}(\lambda,t)=\delta(t-\lambda)$; C = real axis.

(2) $k(t,\lambda)=1(t-\lambda)$ (unit step); $k^{-1}(\lambda,t)=\delta'(\lambda-t)$ (unit doublet function); C = real axis.

(3) $k(t,\lambda)=e^{\lambda t}/2\pi j$; $k^{-1}(\lambda,t)=1(t)e^{-\lambda t}$; C = Bromwich-Wagner contour; $U(\lambda)$ is the Laplace transform of $u(t)$.

(4) $k(t,\lambda)=t^{-\lambda}/2\pi j$; $k^{-1}(\lambda,t)=1(t)t^{\lambda-1}$; C = Bromwich-Wagner contour; $U(\lambda)$ is the Mellin transform of $u(t)$.

(5) $k(t,\lambda)=(\lambda t)^{\frac{1}{2}}J_n(\lambda t)$; $k^{-1}(\lambda,t)=1(t)(\lambda t)^{\frac{1}{2}}J_n(\lambda t)$; C = positive half of the real axis; $U(\lambda)$ is the Hankel transform of order n of $u(t)$.

A linear filter N, whether ideal or not, is characterized analytically by a so-called *characteristic function*[8] $K(t,\lambda)$ which is defined as the response of N to $k(t,\lambda)$—both to be regarded as functions of t involving λ as a parameter. In terms of this function, the relation between the output and input of N may be written as

$$v(t)=\int_C K(t,\lambda)U(\lambda)d\lambda, \tag{16}$$

where $U(\lambda)$ is the spectral function of the input relative to $k(t,\lambda)$. This relation assumes familiar forms when $k(t,\lambda)=\delta(t-\lambda)$ and $k(t,\lambda)=e^{\lambda t}/2\pi j$. In particular, for $k(t,\lambda)=\delta(t-\lambda)$ the corresponding characteristic function $K(t,\lambda)$, that is, the response to $\delta(t-\lambda)$, becomes the *impulsive response* (essentially the one-sided Green's function[9]) of N. In this case $K(t,\lambda)$ is usually written $W(t,\xi)$ with ξ playing the role of λ. Since the spectral function of $u(t)$ relative to $\delta(t-\lambda)$ is equal to $u(\lambda)$, Eq. (16) reduces to the familiar relation

$$v(t)=\int_{-\infty}^{+\infty} W(t,\xi)u(\xi)d\xi. \tag{17}$$

It is to be noted that in the case of a physical system the upper limit in Eq. (17) is t, since the impulsive response of such a system vanishes for $t<\xi$.

We shall need later an expression for the impulsive response of N in terms of a characteristic function of N. This expression can be obtained readily by noting that $K(t,\lambda)$ is the response of N to $k(t,\lambda)$. Thus setting $u(t)=k(t,\lambda)$ in Eq. (17) yields

$$K(t,\lambda)=\int_{-\infty}^{+\infty} W(t,\xi)k(\xi,\lambda)d\xi, \tag{18}$$

and by inversion of this relation we obtain

$$W(t,\xi)=\int_C K(t,\lambda)k^{-1}(\lambda,\xi)d\lambda, \tag{19}$$

which is the desired result.

Turning to ideal filters, consider a set of functions $k(t,\lambda)$ in which the parameter λ belongs to a subset A of C and let $\mathfrak{M}$ be the manifold spanned by this set of functions in Σ. Similarly, let $\mathfrak{N}$ be the manifold spanned by the set $k(t,\lambda)$ in which λ belongs to the complement

[7] See also L. A. Zadeh, Z. angew. Math. Phys. (to be published).

[8] L. A. Zadeh, J. Franklin Inst. (to be published).

[9] K. S. Miller, J. Appl. Phys. **22**, 1054 (1951).

of A with respect to C. Clearly, $\mathfrak{M}$ and $\mathfrak{N}$ constitute a pair of complementary manifolds in Σ and hence can be regarded as the acceptance and rejection manifolds of an ideal filter N. Correspondingly, the definition of an ideal filter may be put into the following form:

A filter N is an ideal linear filter if it passes without distortion all component signals $k(t, \lambda)$ in which the parameter λ belongs to a subset A of the contour C (the pass-band) and rejects all component signals in which λ belongs to the complement of A (the stop-band).

By using this formulation one can obtain at once an expression for the impulsive response of an ideal filter. Thus, from the above definition it follows that the characteristic function of N is of the form

$$\begin{aligned} K(t, \lambda) &= k(t, \lambda) \quad \text{for} \quad \lambda \text{ in } A \\ K(t, \lambda) &= 0 \qquad \text{for} \quad \lambda \text{ not in } A. \end{aligned} \tag{20}$$

Substituting this in Eq. (19), we obtain

$$W(t, \xi) = \int_A k(t, \lambda) k^{-1}(\lambda, \xi) d\lambda, \tag{21}$$

where the integral is taken over the set A.

In general, $W(t, \xi)$ as given by Eq. (21) does not vanish for $t < \xi$ and consequently an ideal filter is not, in general, a physically realizable system. However, as long as $W(t, \xi)$ tends to zero as $t - \xi \rightarrow -\infty$, which it usually does, one can approximately realize $W(t-\beta, \xi)$, where β is a sufficiently large constant with as small an error as desired. In practical terms this means that in general an ideal filter can be approximated to within a constant time-delay by a physical filter.

Two special cases: (a) $k(t, \lambda) = \delta(t-\lambda)$ and (b) $k(t, \lambda) = e^{\lambda t}/2\pi j$ will serve to illustrate the relationships established in the foregoing discussion. For (a) suppose that the set A (that is, the pass-band) consists of the interval $0 \leqq \lambda \leqq T$ on the real axis. Then Eq. (21) yields

$$W(t, \xi) = \int_0^T \delta(t-\lambda)\delta(\lambda-\xi) d\lambda, \tag{22}$$

which reduces to

$$\begin{aligned} W(t, \xi) &= \delta(t-\xi) \quad \text{for} \quad 0 \leqq \xi \leqq T \\ &= 0 \qquad\quad \text{otherwise.} \end{aligned} \tag{23}$$

Thus in this case the ideal filter consists of a switch which is closed during the time interval $0 \leqq t \leqq T$ and is open at other times.

For the second case, (b), let $\lambda = j\omega$ and suppose the set A to consist of the interval $|\omega| < \omega_0$ on the imaginary axis. Since the inverse of $e^{\lambda t}/2\pi j$ is $e^{-\lambda t}$, Eq. (21) yields

$$W(t, \xi) = \frac{1}{2\pi j} \int_{-j\omega_0}^{j\omega_0} e^{j\omega t} e^{-j\omega\xi} d j\omega \tag{24}$$

$$= \frac{\sin\omega_0(t-\xi)}{\pi(t-\xi)}. \tag{25}$$

This represents the impulsive response of a conventional low pass filter. It is observed that $W(t, \xi)$ does not vanish for $t < \xi$ and hence is not physically realizable. However, as was stated previously, one can approximately realize $W(t-\beta, \xi)$ with an arbitrarily small error provided β is sufficiently large.

5. APPLICATIONS

The analogy between ideal filters and projections, and the important role played by the latter in the theory of linear transformations, leads one to believe that ideal filters should eventually occupy an important place in the theory of linear signal transmission systems. In so far as the practical applications of ideal filters are concerned, a preliminary study indicates that such filters might be advantageously used in several applications, particularly in connection with the simultaneous transmission (multiplexing) and filtration of signals. These two applications are briefly discussed below.

Assume for simplicity that two ensembles (sets) of signals $\{u_1(t)\}$ and $\{u_2(t)\}$ are transmitted simultaneously so that at the receiving end one has the sum $\{u_1(t)+u_2(t)\}$. Consider the problem of extracting the $\{u_1(t)\}$ and $\{u_2(t)\}$ ensembles from the sum ensemble $\{u_1(t)+u_2(t)\}$ by means of linear filters. The questions to be answered in this connection are: (a) Under what conditions is it possible to separate $\{u_1(t)\}$ and $\{u_2(t)\}$ by means of such filters? (b) Assuming that these conditions are satisfied, what are the specifications of filters which will perform the separation?

In terms of the notation illustrated in Fig. 1(c), the process of separation of $u_1(t)$ and $u_2(t)$ may be expressed in the following form,

$$(N_1 | N_2)(u_1+u_2) = u_1 | u_2. \tag{26}$$

This is equivalent to the two equations

$$N_1(u_1+u_2) = u_1 \tag{27}$$

and

$$N_2(u_1+u_2) = u_2. \tag{28}$$

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be the manifolds in Σ corresponding to the sets $\{u_1(t)\}$ and $\{u_2(t)\}$, respectively. From Eqs. (27) and (28) it is clear that N_1 and N_2 must be ideal filters such that $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are the acceptance manifolds of N_1 and N_2, respectively. Furthermore, the rejection manifold of N_1 should contain $\mathfrak{M}_2$, while that of N_2 should contain $\mathfrak{M}_1$. Since N_1 and N_2 are assumed to be linear filters, the manifolds $\mathfrak{M}_1$ and $\mathfrak{M}_2$ should be linear and disjoint. Consequently, we can conclude that:

(a) In order to be separable by means of linear filters, the simultaneously transmitted sets of signals $\{u_1(t)\}$ and $\{u_2(t)\}$ should span linear and disjoint manifolds $\mathfrak{M}_1$ and $\mathfrak{M}_2$ in the signal space Σ.

(b) The separation of $\{u_1(t)\}$ and $\{u_2(t)\}$ can be achieved by using two linear ideal filters N_1 and N_2 such that $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are, respectively, the acceptance manifolds of N_1 and N_2. The rejection manifold of N_1

should contain $\mathfrak{M}_2$, while that of N_2 should contain $\mathfrak{M}_1$.

It should be observed that since N_1 and N_2 are realizable, in general, only to within a constant time-delay, one would obtain by separation not $u_1(t)$ and $u_2(t)$ but rather $u_1(t-\beta_1)$ and $u_2(t-\beta_2)$, where β_1 and β_2 are sufficiently large constants.

The conventional methods of multiplexing involve essentially the use of either time or frequency division, both of which are special cases of the more general approach outlined above. Thus, in the case of frequency division, $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are disjoint (and orthogonal) manifolds defined in terms of an exponential basis $e^{\lambda t}/2\pi j$ in the signal space. Similarly, in the case of time division, $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are disjoint (and orthogonal) manifolds defined in terms of a delta-function basis $\delta(t-\lambda)$ in Σ.

The problem of filtration, that is, the problem of extracting the message ensemble $\{\mu(t)\}$ from the sum of $\{\mu(t)\}$ and a noise ensemble $\{\nu(t)\}$ is similar, in principle, to that involved in the separation of two simultaneously transmitted ensembles. One important difference, however, is that whereas in the case of simultaneous transmission of two sets of signals $\{u_1(t)\}$ and $\{u_2(t)\}$ one has complete control over both $\{u_1(t)\}$ and $\{u_2(t)\}$, this is not the case when $\{u_2(t)\}$ represents noise. Thus, in general, one has no control over the manifold $\mathfrak{N}$ corresponding to $\{\nu(t)\}$. If $\mathfrak{N}$ is a linear manifold in Σ, then any suitable complementary manifold $\mathfrak{M}$ in Σ can be assigned to the messages $\{\mu(t)\}$, and in this way one would obtain a communication channel which would not be disturbed (with probability unity) by the noise $\{\nu(t)\}$. If, on the other hand, $\mathfrak{N}$ is not a linear manifold or spans the whole space Σ, then a linear ideal filter cannot be used to separate the message from the noise. However, it may still be possible to achieve a perfect separation of messages and noise by the use of a nonlinear ideal filter. A PCM decoder is an example of such a filter. Other examples of such filters are furnished by the decoding schemes used in the theory of transmission of information.[10,11]

[10] S. O. Rice, Bell Sys. Tech. J. **29**, 60 (1950).
[11] L. A. Zadeh, Trans. N. Y. Acad. Sci. (to be published).

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 2 FEBRUARY, 1952

The Effect of Gas Pressure in the Electric Bulb on the Polarity of the Rectified Current

SHANG-YI CH'EN, *Physics Department, University of Oregon, Eugene, Oregon*

AND

WEI-CHENG LI, *Physics Department, Yenching University, Peking, China*

(Received August 22, 1951)

In sequel to the experiment by Ch'en and Liu (see reference 2) that a rectified current was derivable from the surface of the bulb when the bulb was coated with Chinese ink and was run by alternating current, the following points were observed with 220-v bulbs: (1) When the nitrogen pressure inside the bulb was more than a hundred mm Hg, the polarity of the rectified current was always negative, and the current amounted to a few milliamperes. The current decreased with the decrease of pressure. At a pressure of a few mm Hg the current was decreased to a very small value, and then turned to positive polarity when the pressure was further decreased. The positive current increased to a weak maximum at a few tenths of a mm Hg and came down again as the pressure was further lowered. (2) The current *vs* pressure curves for argon, helium, and neon were essentially similar to those for nitrogen. (3) When the inner surface of the bulb was coated with a layer of metallic sodium, the rectified current, which was positive while without the coating, turned out to be negative. (4) When the material used to paint the bulb contained no sodium compounds, a quick fatigue was observed.

INTRODUCTION

WHEN an electric bulb is connected to a source of alternating current, a rectified current is derivable from the surface of the bulb if the surface is hot and is coated externally by any conducting material. This action was probably first realized by Burt[1] in his experiment on the electrolysis of metallic sodium through glass, and was observed again by Ch'en in an entirely different and much easier way[2] by merely painting the bulb with opaque conducting material. When the bulb was so painted, the bulb wall could be heated by the radiant energy from the incandescent filaments rendering no additional heating necessary as exercised by Burt.

The rectifying characteristics of twenty-three electric bulbs made by eleven different companies were studied by one of the authors and Liu.[2] It was found that the polarity of the painted bulb was negative for some bulbs (hereafter called (−) bulbs), positive for the other (hereafter called (+) bulbs), and for certain bulbs the polarity changed slowly back and forth from negative to positive (hereafter called (±) bulbs). For

[1] R. C. Burt, J. Opt. Soc. Am. **11**, 87 (1925).
[2] S. Y. Ch'en and P. H. Liu, Chinese J. Phys. **7**, 192 (1948).

(−) bulbs metallic sodium was deposited on the inner surface of the bulb, while for (+) bulbs no such deposition took place.

Materials containing no alkali compounds and those with a large percentage of sodium compounds were used to paint the bulbs. So far as the bulb was painted to the same opacity and the paint was a conductor, no appreciable difference as a result of the nature of the paints was observed except the one with no sodium compound which was inclined to show quicker fatigue. Because of the geometrical configuration of the filament[3] with respect to the painted areas on the bulb the amount of derivable current (hereafter called output) varied with different locations on the bulb. The locations which were closest to the filament gave the greatest output. The output ranged from a small fraction of a milliampere to a few milliamperes (for 220-volt 60-w bulbs) depending on the nature of the glass, the thickness and the size of the bulb, the operating voltage, the wattage, and the gas pressure inside the bulb. The output increased with the increase of the temperature of the bulb, reaching a saturated value when the temperature of the bulb was raised to about 110°C for 220-v bulbs. The output increased more than linearly with the increase in the filament voltage. The voltage of the rectified current was about 10–20 volts lower than that of the impressed ac. Both clear and inside-frosted bulbs were used. No difference in behavior was noticed. Bulbs made of soft Na glass exhibited much greater effects than those made of other glass. Pyrex glass gave practically no effect.

Theoretically, the action of the (−) bulb was explained as the result of the neutralization of the sodium ions present in the hot glass by the thermal electrons emitted from the filament. It was the field of the thermal electrons that caused the negative charges in the glass to leave the bulb. The phenomenon for the (+) bulb was explained as due to secondary emission of electrons, which should take place when the pressure inside the bulb was very low. The thermal electrons from the filament could acquire sufficient kinetic energy to knock off electrons on bombarding the glass wall, and at the same time were not caught by the sodium ions of the glass, probably by the help of space charges of positive ions in the bulb. One may consider that there was always a superposition of two sections. In any half-cycle, different areas of the bulb would be doing different things. The area near the negative end of the filament would receive electrons from the filament, and perhaps would have secondary electrons emitted from it. At the same time, areas near the positive end of the filament would lose photoelectrons. It would be a balance among all these things, and perhaps others, if ionization were present, that determined the direction and size of the current observed. (±) bulbs may thus be interpreted as the case where the two actions overbalance each other.

[3] All the bulbs used had a fine coiled tungsten filament mounted roughly in the form of the five sides of a hexagon. The filament was not located accurately at the center of the bulb.

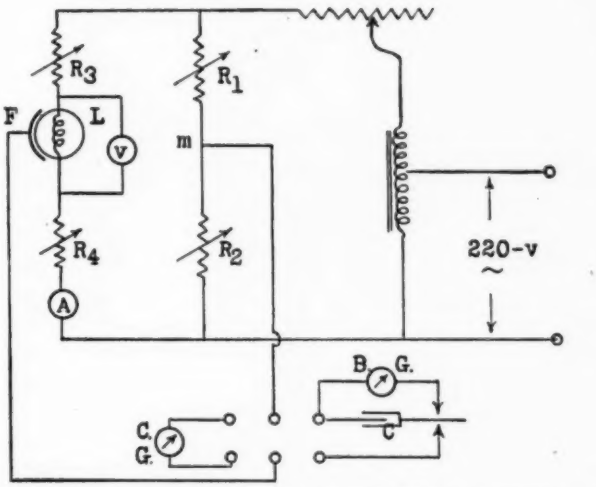


FIG. 1. The circuit. *A*, ammeter; *B. G.*, ballistic galvanometer; *C*, condenser; *C. G.*, current galvanometer; *F*, the coating on the glass bulb; *L*, the bulb to be investigated; $R_1R_2R_3R_4$, variable rheostats; and *V*, the voltmeter.

Gas-filled bulbs (such as G. E. bulbs) were found to behave always as (−) bulbs because the pressure in the bulb was high, and thus the mean free path of the electrons was considerably shortened. The thermal electrons would lose energies by collisions with the gas molecules in the bulb, and would not have sufficient energy of collision with the bulb to produce secondary emission but would be caught by the sodium ions.

It is the purpose of the present article to present further experimental study of (1) the relationship between the polarity of the rectified current and the pressure inside the bulb, and the relationship between the magnitude of the output and the pressure inside the bulb; (2) the effect, if any, of the nature of the gas in the bulb on the output; (3) the effect of the nature of the coating in the interior surface of the bulb on the polarity of the rectified current; and (4) further study on the effect of coating the exterior surface of the bulb with different substances.

EXPERIMENTAL

The circuit used, as shown in Fig. 1, was essentially the same as before,[2] except that variable resistances were used instead of electric bulbs. The potential on the coatings of the bulb was set at a value midway between that of the filament terminals in order to lessen uneven load on the filament resulting from thermal emissions of electrons. The filament carried, in addition to the filament current, also current resulting from the emission of thermal electrons collected by the coated glass bulb. The thermoionic emission would be faster at that portion of the filament whose potential difference to *F* was greater. If the potential at *F* was the same as that of one filament terminal, the filament could be easily burned out near the other filament terminal.

FIG. 2. The bulb with coating and the side tube.

R_1 and R_2 were each 2000 ohms; R_3 and R_4 300 ohms. The potential of F with respect to m, (Fig. 1), was measured with a ballistic galvanometer and its current with a current galvanometer. The temperature of the bulb was measured with a copper-constantan thermocouple. The bulbs used throughout the experiment were G. E. Edison 60-w 220-v bulbs. A glass tubing was connected to the bulb, shown in Fig. 2, so that it could be pumped to any desired vacuum. Four different gases were used to fill the bulb, *viz.*, N_2, A, He, and Ne. The bulbs were coated by Chinese ink which is merely a mixture of finely powdered smoke black, gum, and some common salt.

In the latter part of the experiment the inner surface of the bulb was coated with sodium. The method was to apply the positive terminal of the 110-v dc to the coating and the negative terminal to the filament with a 4000-ohm resistance in series. The entire bulb was painted with carbon black (Chinese ink) except the middle portion which was kept cold by covering it with wet cotton.

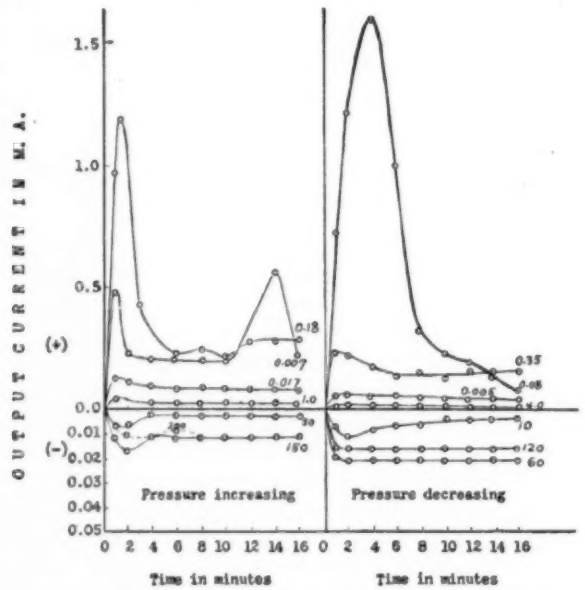


FIG. 3. Rectified current *vs* time curve for Bulb No. 4 filled with N_2.

RESULTS

Several bulbs were studied. Because of variations in geometrical configurations of the filaments with respect to bulb walls, different bulbs usually did not give the same curves, but they all exhibited essentially similar behavior.

Even with the same bulb, the coating of the filament would effect the thermionic emission. Thus different results will be obtained after the filament was heated in the gas. In view of this point observations were made using the same bulb starting with low pressures. Readings were taken each time the pressure was increased to a higher value. Then observations were made by decreasing the pressure.

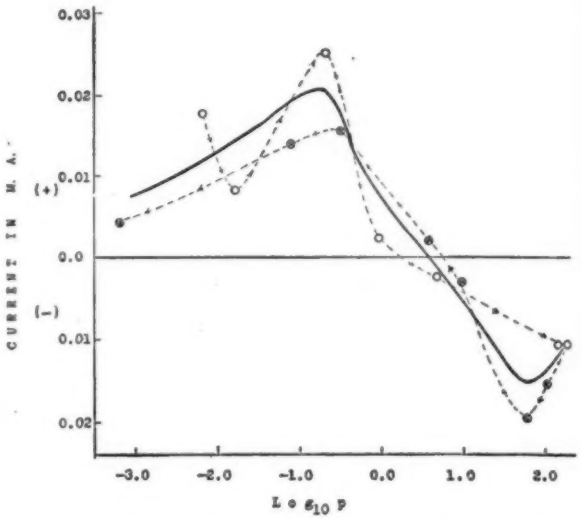


FIG. 4. Steady values of output current *vs* $\log_{10}$ (pressure) for Bulb No. 4. ($p = N_2$ pressure in mm Hg.)

1. The Effect of Nitrogen Pressure in the Bulb on the Output Current

Figure 3 illustrates sample curves obtained with Bulb No. 4, when nitrogen was introduced in the bulb. The numbers marked for each curve were nitrogen pressures in mm Hg. The magnitude of the output current might vary from bulb to bulb, but the current was usually positive when the pressure was below a few mm Hg. This value of the pressure at which the current changes sign would vary with filament voltage, the temperature of the filament, the nature of gas, the distance between the filament and the bulb wall, and probably the concentration of the space charge. This critical pressure for current reversal should not be very different from that at which the mean free path of the

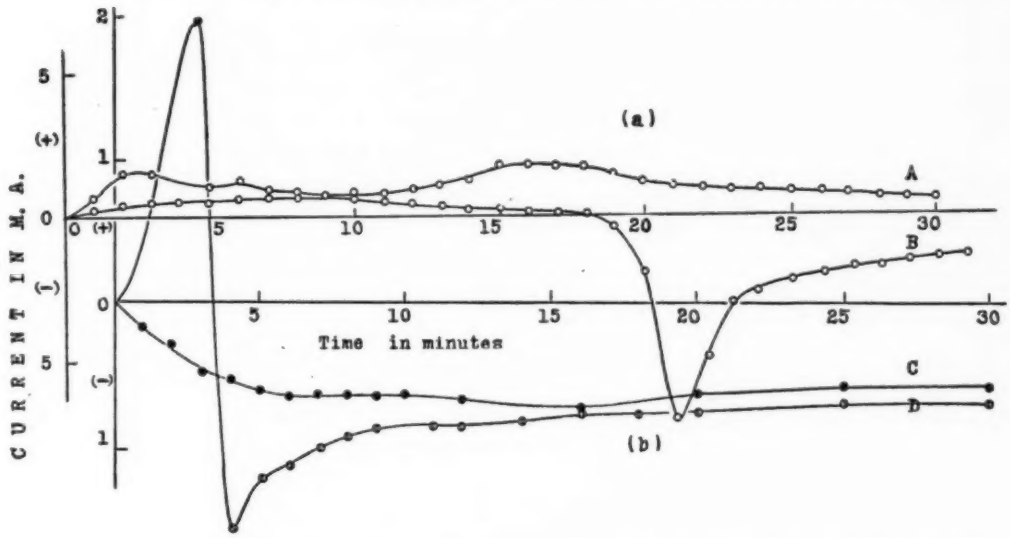


FIG. 5. Some current *vs* time curves for Bulb No. 1. (a) Pressure of N_2 in the bulb = constant = 0.115 mm Hg; imput voltage for Curve A = 180 volts, Curve B = 200 volts. (b) Imput voltage = constant = 200 volts; nitrogen pressure for Curve C = 0.01 mm Hg, Curve D = 0.03 mm Hg.

thermal electrons begins to be comparable with the distance between the filament and the bulb wall. It is not feasible to calculate the theoretical mean free path value accurately because there would be involved many physical quantities which would be very hard to assign, such as the mean temperature of the gas in the bulb, the density and the viscosity of the gas, etc., for the filament temperature was more than one thousand degrees while the temperature of the glass bulb was only of the order of 150°C. The gas pressure and rectified current relationship is shown more clearly in Fig. 4. It is to be noted that the pressure readings were those inside the bulb before the bulb was lighted. Naturally, the pressure, when lighting, was increased by temperature changes and by emitted gas from the filament. Consequently, the low pressure readings were merely semiquantitative.

Among several bulbs tested, all bulbs gave negative rectified current when the pressure was high. This negative current sometimes increased to a weak maximum when the pressure was decreased, say to around 20 mm Hg. As soon as the pressure was decreased still further, the current would decrease to a small value and then start to reverse in direction at a pressure of a few mm of Hg.

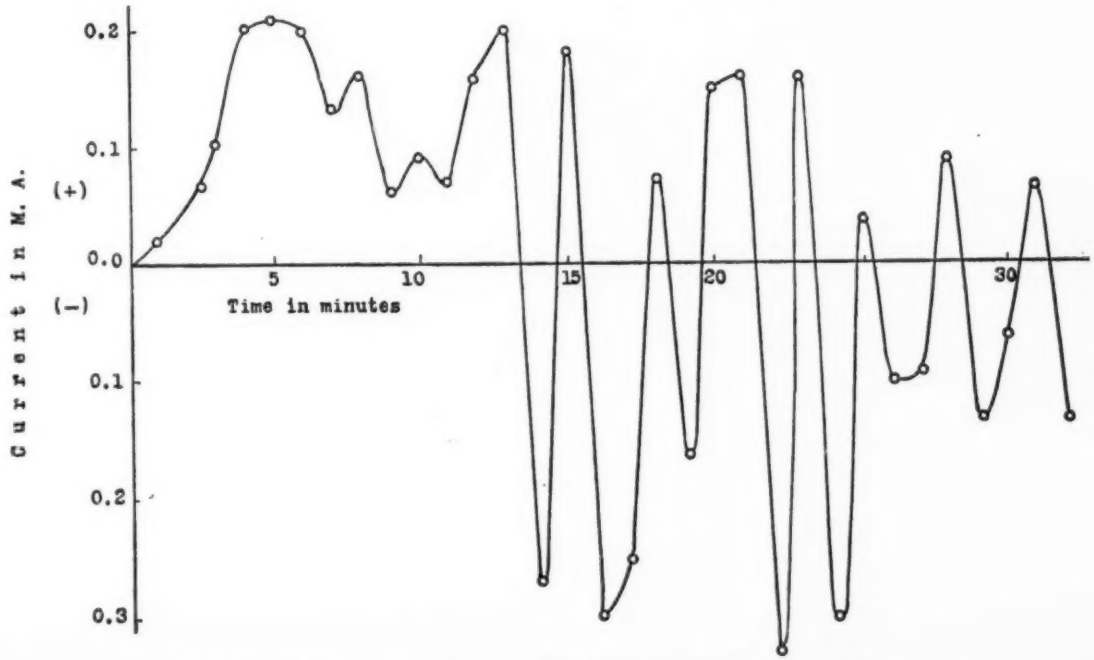


FIG. 6. Rectified current *vs* time curve for Bulb No. 2 filled with N_2, 0.012 mm Hg.

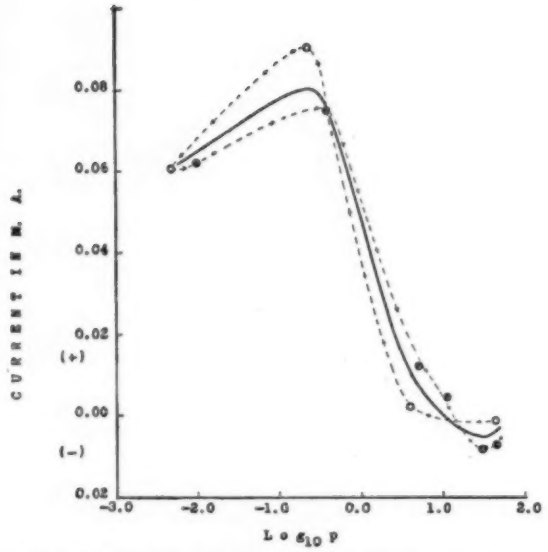


FIG. 7. Steady values of output current *vs* $\log_{10}p$ for Bulb No. 5. (p=argon pressure.)

Different phenomena were observed beyond this pressure point. The bulbs in general showed that the positive current would increase as the pressure was further decreased. As the pressure continued to decrease the current would first increase and then had a tendency to decrease as shown in Fig. 4. But some bulbs first gave a positive current and then the polarity shifted back to negative again as the gas pressure was decreased to the order of 0.1 mm Hg. Below this pressure the current ran continuously negative and had a tendency to go back to positive when the pressure was of the order of 0.001 mm Hg. Phenomena as shown in Figs. 5(b) and 6 would occur at this case.

Another important and interesting result as shown in Fig. 5(a) indicated that the polarity of the rectified current did depend on the imput voltage. The polarity was positive when 180 volts was applied to the bulb as shown in Curve *A*. When 200 volts was applied, the current was first positive and then turned negative as Curve *B*.

Figures 5(b) and 6 show some different phenomena which were observed. In Fig. 5(b) the curve for 0.03 mm pressure first showed a positive current, then it became negative after the bulb was run for a few minutes. Figure 6 exemplified a case that the current shifted back and forth from positive to negative, when the pressure was low.

2. The Output Current under Different Pressures of Other Gases in the Bulb

Similar readings were taken when argon was used to fill the tube. Figure 7 is a typical curve obtained when the pressure of argon in the bulb to start with was 45 mm, and observations were made for pressures down to 0.005 mm and up again, step by step, to 50 mm. The curve gives essentially the same effect for argon as that for nitrogen.

It was very hard to work with helium and neon because it was very easy to discharge through the bulb, thus causing the burning out of the filament. Although no systematic curves for these gases were obtained as for nitrogen and argon, it was quite evident that the results turned out to be essentially like those for the latter gases.

3. The Effect of Coating the Inner Surface of the Bulb with a Layer of Metallic Sodium on the Polarity of the Rectified Current

Two different special bulbs were chosen to test this effect. They were all 220-v 60-w bulbs made in China. The first bulb (Chuen Kwang brand) was a (±) bulb whose output current before and after coating are shown in Fig. 8, lines *A* and *A'*. The second bulb

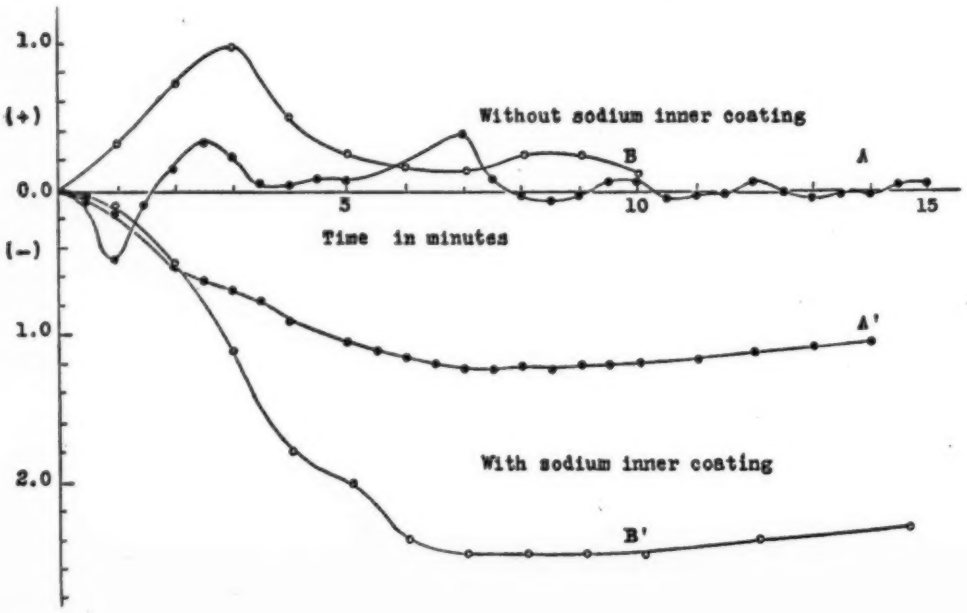


FIG. 8. The effect of coating the inner surface of the bulb with a layer of metallic sodium on the polarity of the rectified current.

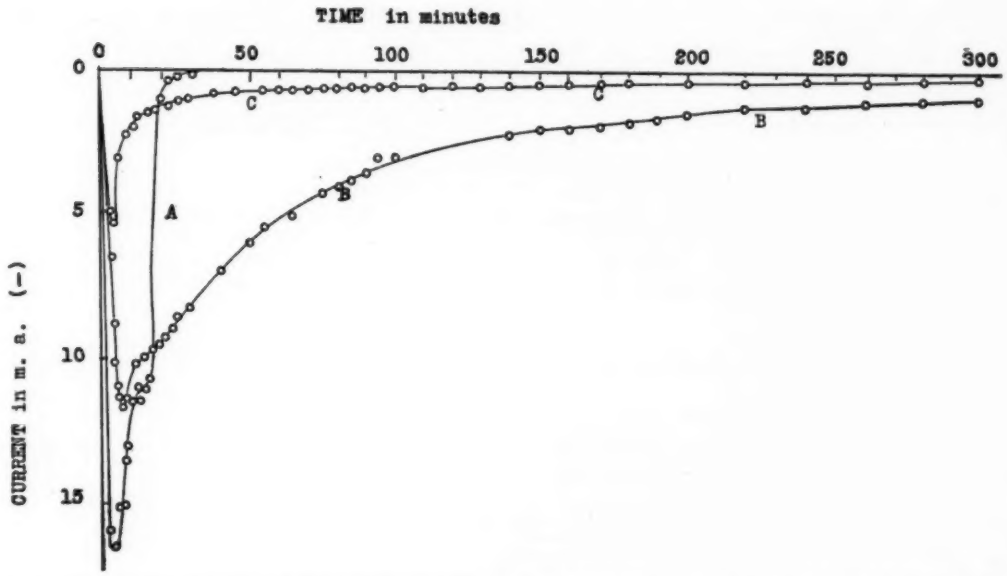


FIG. 9. The effect of coating the exterior surface of the bulb with different substances. Bulb, G. E. Edison; A, silvered surface; B, painted with C and NaCl; C, painted with C only.

(Chi-Te brand) was a (+) bulb which, after coating, behaved as *s* (−) bulb as shown in Fig. 8, lines *B* and *B'*. These curves indicate evidently that coating the bulb wall with sodium tends to destroy the (+) current, which is supposed to be caused by secondary emission and photoelectric effect, so that the collection by the wall of electrons from the filament will be the main process and give a negative current.

4. Further Study on the Effect of Coating the Exterior Surface of the Bulb with Different Substances on the Rectified Current

Figure 9 gives curves obtained when the bulb was coated with silver as the G. E. silvered bowl lamps (Curve C), with a mixture of carbon black and common salt (Curve B), and with pure carbon black and a trace of geleatin (Curve A). After Curve A in Fig. 9 was obtained, the setup was laid untouched for 6 days. When the test was repeated after 6 days one got a curve about the same as before (with steady current values of about 0.3 ma), except that the peak value during the first few minutes had practically disappeared. When Curve B was repeated after 6 days, one got the same thing as that for Curve A. But if one put a few drops of distilled water on the coating and repeated Curve B again, an immediate increase in output current was obtained with initial peak values of about 6 ma, and fatigued down again as Curve B to 2 ma after 60 minutes. It is easy to see that the distilled water would bring more soluble salt in direct contact with the glass bulb.

DISCUSSION

The function of the bulb in these experiments was merely that of a full wave rectifier as recently shown by an oscillograph. The experiments do not lend themselves to involved theoretical analysis, but they suggest some interesting applications. They illustrate well that the electrical properties of hot glass[4] produce appreciable effects on electric bulbs.

The bulb can be used as a very convenient and economical source of rectified dc and gives a steady tone through ear phones. When lights are used in damp places, the life of the bulbs should be influenced by the effect here studied. These effects are more pronounced for 220-v bulbs than for 110-v bulbs. So they should be more important for those who stay in cities supplying only 220-v ac.

For the case such as G. E. silvered bowl lamps, care should be taken not to allow the silver coating to touch the metal fixture if the latter is grounded. The current coming out of the bulb will cause an uneven and additional load on the filament, so the life will be shortened. Also metallic sodium vapor will be produced gradually in the bulb (because gas filled bulbs are always (−) bulbs). The sodium vapor will not only deposit at the cooler parts of the bulb, but will cause the filament to be burned out by a short from the filament to the bulb wall, when the sodium vapor is increased to a certain concentration. In the light of the present experiment, the gas pressure inside the bulb should, in so far as these effects are concerned, be adjusted so that the derivable current from the bulb has a minimum value. For 220-v bulbs the pressure should be around a few mm of Hg. The exact value will depend upon the geometrical as well as the electrical conditions of the bulb. The characteristic properties of power radio tubes should also be effected somewhat by the present effect. The best way to avoid the effect is to use Pyrex glass bulbs.

[4] Such as those mentioned in J. E. Stanworth, *Physical Properties of Glass* (Oxford University Press, London, 1951), Chapter 6.

Considerations Regarding the Distribution of Radiation Damage in Matter Bombarded by Light Ions and a Method of Making it Uniform

PHILIP SCHWED AND GERHART GROETZINGER
Lewis Laboratory, National Advisory Committee for Aeronautics, Cleveland, Ohio
(Received August 16, 1951)

If a beam of high energy particles impinges on a solid, radiation damage is produced only in a very thin layer at the end of the range of the particles. Computations determining the extent of this inhomogeneity of damage are outlined and carried out for the specific case of 20-Mev deuterons impinging on a copper sample. A simple method is discussed for producing uniform damage in a thickness corresponding to the range of the particles.

INTRODUCTION

IN recent years studies of damage in solids due to radiation have been the object of much interest. In the case where the radiation damage produced by monoenergetic ions is the subject of the investigation, there is the difficulty that the damage occurs predominantly near the end of the range of the ions, that is, mainly in a thin layer of the material. Slater in his comprehensive review of this problem[1] states that for the case of protons, e.g., damage occurs primarily after they have slowed down to energies of approximately 10 kev. He concludes, therefore, that the use of cyclotrons for such studies is not only wasteful but also detrimental, as it is difficult to remove the extraneous heat generated by the particles as the slow down to such energies. Furthermore, it is obvious that the damage will be concentrated in a thin layer near the end of the path of the particles and so might have little effect on many bulk properties. These disadvantages can be avoided by the use of low energy particles in conjunction with a thin sample. Such samples are, however, very inconvenient to handle experimentally, and, furthermore, structure sensitive properties show a greater dispersion the thinner the sample. The purpose of this communication is to provide a quantitative estimate of the spatial distribution of the radiation damage produced by monoenergetic ions and to describe a simple method whereby uniform radiation damage can be produced in a relatively thick sample by the use of ions whose range equals the thickness of the sample.

SPATIAL DISTRIBUTION OF THE DAMAGE

No detailed information is available concerning the production of radiation damage by the lightest ions (protons and deuterons). Detailed calculations were done by Ozeroff[2] for the case of fission fragments. His results are, however, not directly applicable to the case of the lightest ions which therefore must be treated separately in the following. In carrying out this computation we will consider separately the region where electronic excitation is dominant and that in which the ion is neutralized.

1. Electronic Excitation Dominant

The energy loss to electronic excitation will be estimated by means of Bethe's basic equation[3] (Born approximation):

$$\frac{dE}{dx}=\frac{-2\pi Z^2 N z^* M e^4}{mE}\log\frac{4Em}{I^*M}, \tag{1}$$

where E is the total energy of the particle, x is the distance it has penetrated into the sample, Z is the atomic number of the particle, N is the number of atoms per unit volume in the sample, z^* is the effective atomic number of the sample (with respect to excitation by ions of energy E), I^* is the effective (geometric) mean excitation energy for electrons of the sample (again with respect to ions of energy E), M is the mass of the incident particle, and m and e are, respectively, the mass and charge of the electron. The essential problem in applying this equation is the determination of good approximate values for Z^* and I^*. For this purpose the range of E here considered is further subdivided into two regions, depending on whether the energy spread of the electrons in a given shell is small or large compared to the average excitation energy of the shell in the question.

For the first case, which occurs for the larger values of E, the following modification of the method of Aron, Hoffman, and Williams[4] is used. At any energy E only those electrons whose effective excitation energies E_i satisfy

$$4Em/ME_i \gtrsim 1 \tag{2}$$

are taken into account so that I^* and Z^* are determined by the equations:

$$z=z^*+n, \tag{3a}$$

$$z\log I=z^*\log I^*+\sum \log E_i, \tag{3b}$$

where z is the atomic number of the material, n is the number of electrons not satisfying (2), and I is the effective mean excitation energy for all the electrons of an atom of the material. The summation in (3b) is to be taken over those electrons *not* satisfying (2). To

[1] J. C. Slater, J. Appl. Phys. **22**, 237 (1951).
[2] J. Ozeroff, KAPL-205 (1949).
[3] M. S. Livingston and H. Bethe, Revs. Modern Phys. **9**, 263 (1937).
[4] Aron, Hoffman, and Williams, AECU-663 (1950).

apply this method it is necessary to know the quantities E_i. We have determined these by multiplying each actual excitation energy[5] by such a factor as to make the mean effective excitation energy equal to that obtained experimentally by Wilson.[6]

For the lower energy range of protons and deuterons the number of electrons with excitation energy greater than a specified amount is estimated by means of the Thomas-Fermi model. The mean excitation energy is taken as equal to the Rydberg energy E_R. The resultant equation in this region is

$$\frac{dE}{dx}=-\frac{2\pi e^4 Z^2 NCE^{\frac{1}{2}}M}{mE}\log\frac{4m}{E_R M}, \tag{4}$$

where C is a constant such that the effective atomic number is given by

$$z^*=CE^{\frac{1}{2}} \tag{5}$$

(see reference 2). The lower energy limit of the region of applicability of this equation is $E_c=(M/m)E_R$ below which the ion will have almost certainly gained an electron so that the energy loss due to electronic excitation diminishes greatly.

The following (classical) expression suffices throughout the range between the initial energy E_0 and the energy E_c to specify the rate of energy loss due to collisions in which an atom of the lattice is displaced (acquires an energy larger than ϵ).

$$-\left(\frac{dE_D}{dx}\right)=\frac{\pi N_0 z^2 Z^2 e^4 M}{M'E}\log\frac{4EM}{\epsilon M'}, \tag{6}$$

where M' is the mass of an atom of the sample. From Eqs. (1) to (6) it is possible to compute E_D', the amount of energy lost in the production of displaced atoms, x' the distance travelled and $(dE_D'/dx)_{Av}$ the average energy loss going into radiation damage per unit length when the ion slows down from E_0 to E_c.

2. Ion Neutralized

At energies below E_c the ion is neutralized and the excitation loss becomes small enough to allow the production of displaced atoms to become essentially an equally important process. We will therefore assume—as Ozeroff[2] did for the case of fission fragments—that the energy loss due to both processes is equal. The range x'' is taken as equal to one-half that which would be obtained solely on the basis of losses due to collisions of the neutralized ion and is given by the following expression:

$$x''=\frac{u^2}{10}\frac{M'}{2\pi N_0 M}\left[E_i\left(2\log\frac{2.2E_c}{\bar{E}}\right)-E_i(2\log 2.2)\right], \tag{7}$$

where $u=z^{\frac{1}{2}}/2a_0$, a_0 is the radius of the Bohr orbit in hydrogen; $\bar{E}=ze^2u$ and

$$E_i(y)=\int_{-\infty}^{y}(e^y dy/y).$$

The average energy loss per unit length $(dE_D''/dx)_{Av}$ in this region is given by

$$\left(\frac{dE_D''}{dx}\right)_{Av}=\frac{E_c}{2x''}. \tag{8}$$

EXAMPLE: DEUTERON BOMBARDMENT OF COPPER

The calculations have been carried out for the case of deuterons of an initial energy $E_0=20$ Mev incident upon copper for which $E_c=50$ kev taking $\epsilon=25$ ev.[2] If the primed and double-primed symbols refer to the status before and after the energy of the particles becomes equal to E_c, the total amount of energy lost in producing displaced atoms is

$$(E_D')_{\text{total}}=8.4\times10^3 \text{ ev}$$
$$(E_D'')_{\text{total}}=2.5\times10^4 \text{ ev};$$

the distance traveled is

$$x'=4.6\times10^{-2} \text{ cm}$$
$$x''=6\times10^{-4} \text{ cm};$$

and the average energy loss per unit path length resulting from the production of displaced atoms becomes

$$(dE_D'/dx)_{Av}=1.8\times10^5 \text{ ev/cm},$$
$$(dE_D''/dx)_{Av}=4.6\times10^7 \text{ ev/cm}.$$

The ratio of the number of displaced atoms per ion produced in the two regions (of thickness x'' and x', respectively) exceeds the value of 200 found for $(E_D'')_{\text{total}}/(E_D')_{\text{total}}$ since the average energy transfer per displaced atom is greater in the region x'. Since the above statements refer only to the rate at which atoms are displaced, the inhomogeneity of the resulting damage will, of course, be diminished if the exposures are extended over a period so long that saturation occurs.

METHODS FOR THE PRODUCTION OF UNIFORM DAMAGE

There are, however, two ways to increase the thickness of a sample beyond x'' without introducing much inhomogeneity prior to saturation of the whole sample. In the first place, an accelerator, such as a cyclotron, with a high energy output and with a large spread of energy, may be used in conjunction with an absorber of an equivalent thickness slightly less than x' placed in front of the sample. The spread of energies present in the original beam combines with the effects of straggling and scattering introduced by the absorber to in-

[5] The International Critical Tables.
[6] R. R. Wilson, Phys. Rev. **60**, 749 (1941).

crease the thickness within which the rate of production of displaced atoms is reasonably uniform. However, a considerably greater improvement can be obtained by the use of a device which introduces an absorber into the beam in such a fashion that the thickness of material traversed by the particles varies linearly and periodically between zero and the range of the particles in it. With this device it becomes possible to use a sample of thickness $x'+x''$ since the region in which the particles have energies less than E_e moves through the whole thickness of the sample in such a way that the rate of production of displaced atoms is essentially uniform. The heat generated in the sample is reduced by a factor of about two and the induced radioactivity by a much larger factor compared to that produced without such an absorber.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 2 FEBRUARY, 1952

Studies in Newtonian Flow. III. The Dependence of the Viscosity of Liquids on Molecular Weight and Free Space (in Homologous Series)

ARTHUR K. DOOLITTLE
Carbide and Carbon Chemicals Company, A Division of Union Carbide and Carbon Corporation, South Charleston, West Virginia
(Received September 5, 1951)

The free-space concept, previously applied to variation resulting from a change in temperature only, is here adapted to a case where both temperature and molecular weight vary. The molecular weight range of the *n*-paraffins illustrative of this case is limited to $m=100$ through $m=240$. Over this range of molecular weights the family of lines represented by $\ln\eta=B(v_0/v_f)+\ln A$ intersects the vertical axis at very nearly a common point. By assuming a common intercept and representing the slopes of this family of lines in terms of molecular weight, an expression defining viscosity as a function of molecular weight and free space is deduced. This expression reproduces the "selected data" satisfactorily over the molecular weight range mentioned. Its greater significance as a step in the development of a far more useful function will become apparent in the succeeding paper.

THE dependence of the viscosity of liquids on molecular weight, as well as its dependence on temperature and on free space which were discussed in the two prior papers of this series,[1,2] have been studied by many investigators over a period of years. Usually, though not always, such studies have dealt with homologous series of liquids. A resumé of a number of expressions relating viscosity to molecular weight may be found in a chapter by the author in Alexander's *Colloid Chemistry*.[3]

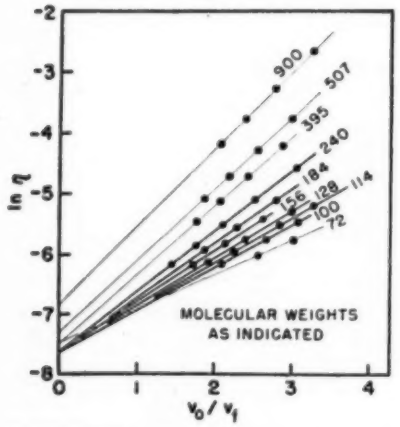


FIG. 1. *n*-Paraffin hydrocarbons. Logarithm of viscosity *vs* reciprocal relative free space. Plot of "selected data," all molecular weights, showing convergence of lines to a common intercept for molecular weights $m=100$ through $m=240$.

The study described in this paper is based on accurate measurements of the density and viscosity of several members of the homologous series of *n*-alkanes. The synthesis and measurement of these compounds has been described in a prior publication.[4]

In the previous paper of this series[2] it was shown that for the *n*-paraffins the variation in viscosity due to a change in temperature only (except for those values in the neighborhood of their freezing points) could be expressed accurately as a function of free space by the following equation,

$$\eta=Ae^{B/(v_0/v_f)} \quad \text{or} \quad \ln\eta=B(v_0/v_f)+\ln A, \qquad (1)$$

where η = coefficient of viscosity, v_f/v_0 = relative free-space for a single substance at fixed pressure, and A and B = constants for a single substance.

We now undertake to show how the above relationship may be extended from the consideration of a single compound at a time to the treatment of members of an homologous series of compounds. It should be emphasized, however, that whereas Eq. (1) may be an exact function for temperatures well above the freezing point, and certainly has been shown to be as precise as the accuracy of the measurements, no such claim is made for the molecular weight function to be developed in this paper. The reasons for the slight inaccuracies in-

[1] A. K. Doolittle, J. Appl. Phys. **22**, 1031 (1951).
[2] A. K. Doolittle, J. Appl. Phys. **22**, 1471 (1951).
[3] J. Alexander, *Colloid Chemistry* (Reinhold Publishing Corporation, New York, 1950), Vol. VII, pp. 155–156.
[4] A. K. Doolittle and R. H. Peterson, J. Am. Chem. Soc. **73**, 2145 (1951).

troduced in deriving the molecular weight function will become apparent shortly.

In proceeding to adapt Eq. (1) to the more general case of the dependence of viscosity on relative free space, when the latter varies as a result of a change in either temperature or molecular weight or both simultaneously, let us first consider the family of lines obtained by plotting the "selected data"* for all members of the homologous series of *n*-alkanes used in our study.[4] These data are plotted according to the logarithmic form of Eq. (1) and are shown on Fig. 1 (from Table I). The slopes and intercepts of those lines containing four or more valid points were calculated by the least squares method mentioned in the previous paper.[2] The slopes of lines containing less than four valid points were approximated graphically. In no case, however, does the "selected measured value" deviate from the chosen line by more than the accuracy of the measurement.

It will be observed that all of the straight lines defined by the "selected values" for paraffins $m=100$ through $m=240$ intersect the vertical axis at approximately the same point. The lines defined by molecular weights below $m=100$ and above $m=240$ intersect the vertical axis above this common point. The relationship that will shortly be developed will be limited, therefore, to *n*-paraffins of molecular weight 100 through 240.

For the purpose of deriving a relationship to represent the case where either the temperature or the molecular weight (or both) vary, we arbitrarily make the assumption that the lines representing molecular weights 100 through 240 do intersect at a common point (0, −7.6500) which is approximately the average of their actual individual intercepts shown in Table I. The slight adjustments made by employing the average intercept necessitates a corresponding adjustment in the slopes so that the values we shall now use will be very slightly different from those of Table I which were determined from least squares and graphical calculations. This, then, introduces an inaccuracy that detracts from the absolute validity of this free space—molecular weight function.

The next step in our development will be more readily understood if we first consider briefly certain aspects of the nature of viscosity. Viscosity is the resistance that must be overcome when a liquid is continuously deformed under stress. This resistance arises from the necessity of moving the molecules of the liquid past each other against the cohesive forces tending to hold them together. For a given substance, the intensity of the force field at unit distance from any given molecule is presumed to be fixed, the change in internal pressure with the temperature being occasioned by a change in the average separation of the molecules from each other. Since Eq. (1) relates the resistance to flow to a constant, B, divided by the separation of the molecules, (v_f/v_0), it implies that B should be some measure of the

* The "selected data" are data from which values in the neighborhood of the freezing points have been excluded, see reference 2.

TABLE I. *n*-Paraffin hydrocarbons. Reciprocal relative free space function; comparison of viscosity values determined by best straight lines with corresponding measured values, $\ln\eta_{calc}=B(v_0/v_f)+\ln A$.

$m\ \uparrow\uparrow\uparrow$ $B\ \uparrow\uparrow\uparrow$ $\ln A\ \uparrow\uparrow\uparrow$					72.15 0.5550 −7.4552				100.20 0.6880 −7.6070				114.22 0.7400 −7.6446				128.25 0.7780 −7.6472			
T	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	% Acc	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	% Acc	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	% Acc	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	% Acc
263.2	−5.7708	0.003117	0.003114	+0.1	0.4															
293.2	−6.0322	0.002400	0.002404	−0.2	0.4	−5.4790	0.004174	0.004180	−0.1	0.2	−5.2093	0.005466	0.005479	−0.2	0.2					
323.2						−5.7721	0.003113	0.003100	+0.4	0.4	−5.5363	0.003941	0.003928	+0.3	0.3	−5.3070	0.004957	0.004968	−0.2	0.3
373.2						−6.1705	0.002090	0.002099	−0.4	0.5	−5.9729	0.002547	0.002554	−0.3	0.5	−5.7777	0.003096	0.003087	+0.3	0.4
423.2																−6.1503	0.002133	0.002136	−0.1	0.5

$m\ \uparrow\uparrow\uparrow$ $B\ \uparrow\uparrow\uparrow$ $\ln A\ \uparrow\uparrow\uparrow$					156.30 0.8515 −7.6705				184.35 0.9119 −7.6785				240.46 0.9995 −7.6505				394.74 1.1289 −7.4731			
T	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	% Acc	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	% Acc	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	% Acc	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	% Acc
373.2	−5.4268	0.004397	0.004384	+0.3	0.3	−5.1170	0.005994	0.006001	−0.1	0.2	−4.5880	0.01017	0.01018	−0.1	0.2					
423.2	−5.8483	0.002885	0.002894	−0.3	0.5	−5.5784	0.003779	0.003765	+0.4	0.4	−5.1158	0.006001	0.005984	+0.3	0.3	−4.2253	0.01462	0.01460	+0.1	1.3
473.2	−6.1957	0.002038	0.002048	−0.5	0.6	−5.9495	0.002607	0.002615	−0.3	0.6	−5.5376	0.003936	0.003938	−0.1	0.5	−4.7265	0.008858	0.008856	0.0	1.2
523.2											−5.8894	0.002769	0.002785	−0.6	0.6	−5.1318	0.005906	0.005920	−0.2	1.3
573.2											−6.1952	0.002039	0.002030	+0.4	0.8	−5.4783	0.004177	0.004173	+0.1	1.5

$m\ \uparrow\uparrow$ $B\ \uparrow\uparrow$ $\ln A\ \uparrow\uparrow$					506.95 1.1773 −7.3137				899.68 1.2853 −6.8622					
T	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	% Acc	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	% Acc	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	% Acc
423.2	−3.7747	0.02295	0.02315	−0.9	1.4	−2.6464	0.07091	0.07144	−0.7	1.5					
473.2	−4.3033	0.01352	0.01344	+0.6	1.4	−3.2711	0.03797	0.03757	+1.1	1.6					
523.2	−4.7366	0.008769	0.008749	+0.2	1.4	−3.7775	0.02288	0.02292	−0.2	1.5					
573.2	−5.1016	0.006087	0.006083	+0.1	1.5	−4.1875	0.01519	0.01525	−0.4	1.5					

TABLE II. *n*-Paraffin hydrocarbons. Calculation of $C \cdot f(m) = \dfrac{\ln\eta + |\ln A|_{100-240}}{v_0/v_f}$, $|\ln A|_{100-240} = 7.6500$.

$m \rightarrow$	100.20				114.22				128.25				156.30				184.35				240.46			
T	v_0/v_f	$\ln\eta$	+7.6500	$C \cdot f(m)$	v_0/v_f	$\ln\eta$	+7.6500	$C \cdot f(m)$	v_0/v_f	$\ln\eta$	+7.6500	$C \cdot f(m)$	v_0/v_f	$\ln\eta$	+7.6500	$C \cdot f(m)$	v_0/v_f	$\ln\eta$	+7.6500	$C \cdot f(m)$	v_0/v_f	$\ln\eta$	+7.6500	$C \cdot f(m)$
293.2	3.093	−5.477	2.1730	0.7026	3.291	−5.207	2.4430	0.7423																
323.2	2.667	−5.776	1.8740	0.7027	2.849	−5.540	2.1100	0.7406	3.008	−5.305	2.3450	0.7796												
373.2	2.088	−6.166	1.4840	0.7107	2.259	−5.970	1.6800	0.7437	2.403	−5.781	1.8690	0.7778	2.635	−5.430	2.2200	0.8425	2.809	−5.116	2.5340	0.9021	3.064	−4.587	3.0630	0.9997
423.2									1.924	−6.149	1.5010	0.7801	2.140	−5.845	1.8050	0.8435	2.303	−5.582	2.0680	0.8980	2.536	−5.119	2.5310	0.9980
473.2													1.732	−6.191	1.4590	0.8424	1.896	−5.947	1.7030	0.8982	2.114	−5.537	2.1130	0.9995
523.2																					1.762	−5.884	1.7660	1.0023
573.2																					1.456	−6.200	1.4500	0.9959
Average				0.7053				0.7422				0.7792				0.8428				0.8994				0.9991

TABLE III. *n*-Paraffin hydrocarbons. Dependence of viscosity on reciprocal relative free space when both temperature and molecular weight vary; $\ln\eta_{calc} = 2401 e^{-10.229/m^{1/20}} \cdot (v_0/v_f) - 7.6500$.

$m \rightarrow$			100.20					114.22					128.25				
$Ce^{-K/m^{1/20}} \rightarrow$			0.7056					0.7440					0.7796				
$(=D)$																	
T	v_0/v_f	$D \cdot (v_0/v_f)$	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	v_0/v_f	$D \cdot (v_0/v_f)$	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	v_0/v_f	$D \cdot (v_0/v_f)$	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev
293.2	3.093	2.1824	−5.4676	0.004222	0.004180	+1.0	3.291	2.4485	−5.2015	0.005508	0.005479	+0.5						
323.2	2.667	1.8818	−5.7682	0.003126	0.003100	+0.8	2.849	2.1197	−5.5303	0.003965	0.003928	+0.9	3.008	2.3450	−5.3050	0.004967	0.004968	0.0
373.2	2.088	1.4733	−6.1767	0.002077	0.002099	+1.0	2.259	1.6807	−5.9693	0.002556	0.002554	+0.1	2.403	1.8734	−5.7766	0.003099	0.003087	+0.4
423.2													1.924	1.5000	−6.1500	0.002134	0.002136	−0.1

$m \rightarrow$			156.30					184.35					240.46				
$Ce^{-K/m^{1/20}} \rightarrow$			0.8437					0.9008					0.9995				
$(=D)$																	
T	v_0/v_f	$D \cdot (v_0/v_f)$	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	v_0/v_f	$D \cdot (v_0/v_f)$	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev	v_0/v_f	$D \cdot (v_0/v_f)$	$\ln\eta_{calc}$	η_{calc}	η_{obs}	% Dev
373.2	2.635	2.2231	−5.4269	0.004397	0.004384	+0.3	2.809	2.5303	−5.1197	0.005978	0.006001	−0.4	3.064	3.0625	−4.5875	0.01018	0.01018	0.0
423.2	2.140	1.8055	−5.8445	0.002896	0.002894	+0.1	2.303	2.0745	−5.5755	0.003790	0.003765	+0.7	2.536	2.5347	−5.1153	0.006004	0.005984	+0.3
473.2	1.732	1.4613	−6.1887	0.002054	0.002048	+0.3	1.896	1.7079	−5.9421	0.002627	0.002615	+0.5	2.114	2.1129	−5.5371	0.003938	0.003938	0.0
523.2													1.762	1.7611	−5.8889	0.002770	0.002785	−0.5
573.2													1.456	1.4553	−6.1947	0.002040	0.002030	+0.5

force field intensity within the liquid. Equation (1) also indicates that the change in the magnitude of the cohesive forces in the liquid is occasioned solely by the change in v_f/v_0, or in other words, by the extent of separation of the molecules from each other. If, however, we vary the molecular weight, we alter not only the internal attraction by changing the distances that the molecules are from each other, but we alter the force field intensity as well.

By adjusting our point of view in regard to Eq. (1) in conformance with the foregoing physical picture, we may extend the scope of this equation to permit its application to the case where the molecular weight (in homologous series), as well as the temperature, is varied. Thus, since $\ln A$ is assumed to be invariant, $1/(v_f/v_0)$ or v_0/v_f may be thought of as approaching zero by either of two routes, the value of the limit being the same by either route.

Accepting the value of the common intercept as the average of the intercepts of the individual lines ($m=100$ through $m=240$), we now proceed to determine the slopes that best represent the family of lines having this common intercept. There are, no doubt, several ways in which this might be done, but the plan adopted by us was to calculate the slopes of the lines determined by each point and the common intercept, then average these slopes. Since each slope is some function of molecular weight, we write

$$\text{slope}=Cf(m)=(\ln\eta+7.65/v_0/v_f). \qquad (2)$$

Calculation of $Cf(m)$ for our "selected values" ($m=100$ through $m=240$) was made. The results are presented in Table II where the average values of $Cf(m)$ are shown at the bottom of the corresponding columns. From the form of the curve $Cf(m)$ *vs* m, it was established that the relationship applying in the range $m=100$ to $m=240$ could be represented by a logarithmic decrement in which the argument varies as $1/m^{1/20}$. That is,

$$f(m)=\exp(-K/m^{1/20}) \quad \text{and}$$
$$\ln[C(fm)]=\ln C-K/m^{1/20}. \qquad (3)$$

The final step leading to the molecular weight function involves merely substituting

$$Cf(m)=C\exp(-K/m^{1/20})$$

for B in Eq. (1), giving

$$\ln\eta=C\,[\exp(-K/m^{1/20})](v_0/v_f)+\ln A. \qquad (4)$$

Taking the values for our n-paraffins determined as described above, Eq. (4) becomes

$$\ln\eta_{\text{calc}}=2401\,[\exp(-10.239/m^{1/20})](v_0/v_f)-7.6500. \qquad (5)$$

Values of η_{calc} calculated according to Eq. (5) for our "selected data" in the molecular weight range 100 through 240 are shown in Table III. No value differs from the measured value by more than 1 percent. This equation is neither valid for paraffins of molecular weight less than 100 nor greater than 240.

Although an equation that applies only to the n-alkanes and that is not even valid for all of the members of this series may seem to be rather impractical, the development of this equation is, nevertheless, a necessary step in the sequence of relationships presented in this series of papers. The succeeding and final paper of our "Studies in Newtonian Flow" adapts Eq. (4) to the more general case of the dependence of viscosity on molecular weight (at constant temperature) in any homologous series of Newtonian liquids. The latter relationship is of quite general usefulness.

SUMMARY

(a) The intercepts of the lines represented by $\ln\eta=B(v_0/v_f)+\ln A$ for n-paraffins of molecular weight $m=100$ through $m=240$ intersect the vertical axis at very nearly a common point.

(b) The slopes of the family of lines having a common intercept (n-paraffins $m=100$ through $m=240$) were found to conform to a logarithmic decrement relationship with respect to molecular weight in which the argument varies as $1/m^{1/20}$.

(c) A molecular weight—free space function applying to n-paraffins $m=100$ through $m=240$ was derived from the free space equation, $\ln\eta=B(v_0/v_f)+\ln A$, by substituting for B its equivalent in terms of molecular weight. The molecular weight—free space equation was thus shown to be:

$$\ln\eta=C\,[\exp(-K/m^{1/20})](v_0/v_f)+\ln A.$$

This equation is not an exact function, but it does reproduce the "selected measured values" with an accuracy of 1 percent or better.

Space-Charge-Limited Currents between Inclined Plane Electrodes

HENRY F. IVEY
Research Department, Westinghouse Electric Corporation, Bloomfield, New Jersey
(Received August 23, 1951)

The method of attack devised by Walker has been applied to the problem of current flow between inclined plane electrodes. It is assumed that the initial velocities of emission are negligible. Solutions for the potential distribution, the space-charge characteristic ("perveance"), the particle trajectories, and the transit time in such diodes under conditions of complete space-charge limitation are given. The trajectories and transit time have also been determined for the space-charge-free case. The results are presented in both graphical and tabular form. In most cases the solutions are given for angles between the electrodes as great as 180°.

For an angle between the electrodes of approximately 118° the transit time in the space-charge-free case is equal to that for complete space charge limitation. The effect of accidental electrode tilt on the perveance of a diode, which presumably has parallel electrodes, has also been investigated. A rectangular electrode whose width is 1.24 times the interelectrode spacing is relatively insensitive to tilt about a line passing through its center and parallel to its side. A disk-shaped electrode is insensitive to tilt about any axis passing through its center if its diameter is 1.42 times the interelectrode spacing.

1. INTRODUCTION

THE solutions of the equations for space-charge-limited flow between parallel planes,[1,2] concentric cylinders,[2,3] and concentric spheres[4] have been known for a long time. These three cases are all characterized by rectilinear flow, i.e., the electron (or ion) trajectories are straight lines (assuming that initial velocities of emission are negligible or are normal to the emitter surface). Recently attempts have been made to solve the space charge equations for more general cases involving curvilinear flow.[5–7]

Walker[6] has applied the principles of hydrodynamics to space-charge-limited currents with considerable success. The discussion is limited to "congruent" or irrotational flow, i.e., to cases where

$$\text{curl}\mathbf{v}=0, \tag{1}$$

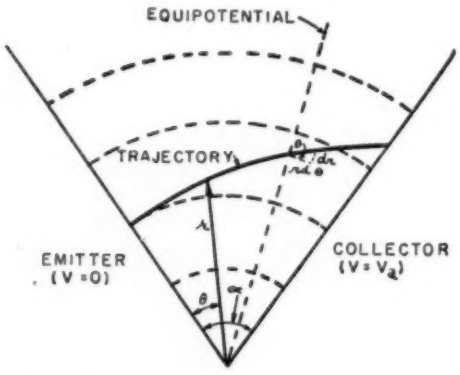


FIG. 1. Electrode arrangement and coordinate system employed. (The angle between the trajectory and the normal to the equipotential should be labeled ϕ, not θ.)

where $\mathbf{v}$ is the velocity vector. This restriction implies that a single trajectory passes through any given point in space and that the velocity is a single valued function of position. It may be stated here that situations involving initial velocities of emission or nonequipotential emitters violate this condition.

Walker[6] has, in particular, shown that the case of inclined plane electrodes is susceptible of solution and has roughly outlined this solution. It is the purpose of the present paper to present in detail the results for this electrode configuration, following his method, and extending it in some cases.

2. THE FUNDAMENTAL EQUATIONS

For irrotational flow the velocity may be expressed in terms of a scalar point function, the action function W, as

$$\mathbf{v}=\text{grad}W. \tag{2}$$

Application of the force equation (Newton's second law of motion) leads to

$$(\text{grad}W)^2=(2e/m)V+\text{const.} \tag{3}$$

If the velocity is zero when $V=0$, then the constant may also be set equal to zero. Equation (3) simply expresses the conservation of energy.

Consider now the electrode arrangement shown in Fig. 1. It is assumed that the length of the electrodes normal to the plane of the figure is great enough (or that end plates are provided) so that end effects are negligible and thus there will be no component of flow perpendicular to this plane. The trajectories can then be described by the polar coordinates r and θ. Following Walker, let us assume an action function of the form

$$W=rg(\theta), \tag{4}$$

so that the variables are separated. This assumption will be seen later to be justified since it leads to solutions of the differential equations which fit the boundary

[1] C. D. Child, Phys. Rev. **32**, 492 (1911).
[2] I. Langmuir, Phys. Rev. **2**, 450 (1913).
[3] I. Langmuir and K. Blodgett, Phys. Rev. **22**, 347 (1923).
[4] I. Langmuir and K. Blodgett, Phys. Rev. **24**, 49 (1924).
[5] B. Meltzer, Proc. Phys. Soc. (London) **62B**, 431, 813 (1949).
[6] G. B. Walker, Proc. Phys. Soc. (London) **63B**, 1017 (1950).
[7] K. Spangenberg, J. Franklin Inst. **232**, 365 (1941).

conditions. The components of the gradient in polar coordinates are

$$v_r=\dot{r}=\text{grad}_r W=\frac{\partial W}{\partial r}=g, \tag{5a}$$

$$v_\theta=r\dot{\theta}=\text{grad}_\theta W=\frac{1}{r}\frac{\partial W}{\partial \theta}=\frac{dg}{d\theta}=g'. \tag{5b}$$

Equation (3) thus becomes

$$g^2+g'^2=CV, \tag{6}$$

where

$$C=2e/m. \tag{7}$$

Equation (6) implies that the potential is dependent on the angle variable θ only, i.e., the equipotentials are radial lines, $\theta=$constant.

Poisson's equation and the equation of continuity remain to be satisfied. Since V is independent of r, the former may be written

$$V''/r^2=4\pi\rho, \tag{8}$$

where ρ is the magnitude of the (negative) space-charge density. The equation of continuity, $\text{div}\mathbf{j}=0$, becomes in polar coordinates

$$\frac{1}{r}\frac{\partial}{\partial r}(r\rho v_r)+\frac{1}{r}\frac{\partial}{\partial \theta}(\rho v_\theta)=0. \tag{9}$$

When Eqs. (5a), (5b), and (8) are substituted, the resulting expression is

$$\frac{d}{d\theta}(V''g')-V''g=0. \tag{10}$$

The two Eqs. (6) and (10), when solved simultaneously and subject to the appropriate boundary conditions, give the desired solutions for the potential and velocity functions.

The trajectories for this flow can be obtained very simply from Eqs. (5a) and (5b).

$$\frac{g}{g'}=\frac{\dot{r}}{r\dot{\theta}}=\frac{dr}{rd\theta} \tag{11}$$

or

$$r=r_0\exp\left(\int_0^\theta \frac{g}{g'}d\theta\right)=r_0h(\theta), \tag{12}$$

where r_0 is the value of r for $\theta=0$. It is seen that the ratio r/r_0 is a function of θ only. This means that the trajectories all have the same "magnification factor," as pointed out by Walker. An example of a trajectory is shown in Fig. 1.

The angle between the trajectories and the equipotentials is also of interest. Let ϕ be defined as the angle between the trajectory and the normal to the equipotential (see Fig. 1). Then

$$\tan\varphi=\frac{dr}{rd\theta}=\frac{g}{g'}. \tag{13}$$

The transit time may be obtained from the following considerations. From (5b)

$$dt=(r/g')d\theta,$$

so that for $t=0$ when $\theta=0$,

$$t=\int_0^\theta \frac{r}{g'}d\theta=r_0\int_0^\theta \frac{h}{g'}d\theta, \tag{14}$$

where h is given by Eq. (12). Since h and g' are functions of θ only, it is seen that the transit time is directly proportional to r_0, the initial value of r.

Let us now consider the electrical current density resulting from the flow. The direction of the current with respect to the equipotentials (and therefore with respect to the electrodes) is given by the angle ϕ of Eq. (13). The velocity is given by the energy equation and ρ is given by Eq. (8), so that the current density at an electrode is

$$J=|\rho\mathbf{v}|\cos\varphi=\frac{1}{4\pi}\left(\frac{2e}{m}\right)^{\frac{1}{2}}\frac{V''V^{\frac{1}{2}}}{r^2}\cos\varphi. \tag{15}$$

This is the analog to the usual space-charge equation.

In the following sections, the solutions for the completely space charge limited and space-charge-free cases will be discussed. As yet the solution for partial space-charge limitation (i.e., electric field at the emitter not zero) has not been found.

3. SOLUTIONS FOR COMPLETE SPACE-CHARGE LIMITATION

Walker has shown that the solutions for the case of completely space-charge-limited flow may be written in series form as

$$g=\sum_{n=1}^{\infty}\alpha_n\theta^{(6n-1)/3}, \tag{16}$$

and

$$V=\sum_{n=1}^{\infty}\beta_n\theta^{(6n-2)/3}=\beta_1\theta^{4/3}f(\theta)=\beta_1\tilde{f}(\theta), \tag{17}$$

where

$$f(\theta)=\sum_{n=0}^{\infty}b_n\theta^{2n}. \quad (b_0=1) \tag{17a}$$

That these solutions do correspond to complete space-charge limitation (i.e., zero electric field at the emitter) can be seen by noting that V does not contain a linear term in θ, and hence the cathode field, which is proportional to V', does indeed disappear for $\theta=0$.

The relations between the α_n and b_n coefficients are found by substituting the series given by expressions

TABLE I. Values of the coefficients for the various series.

n	α_n/α_1	a_n	b_n	c_n/c_1	d_n	p_n	q_n
0	...	...	+1.0000000	...	...	+1.0000000	+1.0000000
1	+1.0000000	+0.3000000	$+2.6666666\times10^{-2}$	+1.0000000	+0.3333333	$+6.6666666\times10^{-2}$	-0.2222222
2	-7.5757575×10^{-2}	$+1.36363636\times10^{-3}$	$+2.3434343\times10^{-3}$	-9.5238095×10^{-2}	$+2.1164021\times10^{-2}$	$+9.3737374\times10^{-3}$	$+1.7636684\times10^{-3}$
3	$+4.2810457\times10^{-3}$	$+4.8770005\times10^{-4}$	$+2.166104\times10^{-4}$	$+4.8100048\times10^{-3}$	$+1.7102239\times10^{-3}$	$+1.1913570\times10^{-3}$	-2.3413778×10^{-4}
4	-8.69635×10^{-4}	-3.593185×10^{-5}	$+1.803295\times10^{-5}$	-4.6827556×10^{-5}	$+1.1401492\times10^{-4}$	$+1.262306\times10^{-4}$	$+1.685340\times10^{-5}$
5	$+2.81097\times10^{-4}$	-9.64152×10^{-6}	$+1.18440\times10^{-6}$	$+2.661063\times10^{-6}$	$+3.207674\times10^{-6}$	$+1.00674\times10^{-5}$	-4.315653×10^{-7}
6	-9.37580×10^{-5}	-1.02665×10^{-6}	$+3.78783\times10^{-8}$	-5.629113×10^{-8}	-6.443824×10^{-7}	$+3.80065\times10^{-7}$	-1.827888×10^{-7}
7	-2.49678×10^{-5}	-4.0391×10^{-8}	-4.20727×10^{-9}	-2.030987×10^{-8}	-1.499003×10^{-7}	-4.81524×10^{-8}	-2.859004×10^{-8}
8	-2.5464×10^{-6}	$+6.9892\times10^{-9}$	-9.49476×10^{-10}	-2.766778×10^{-9}	-1.862019×10^{-8}	-1.23184×10^{-8}	
9	-1.1833×10^{-10}	$+1.6858\times10^{-9}$					

(16) and (17) into Eqs. (6) and (10). The result from (6) is

$$9C\beta_1 b_{n-1}=9\sum_{j+k=n}\alpha_j\alpha_k + \sum_{j+k=n+1}(6j-1)(6k-1)\alpha_j\alpha_k, \quad (n\geqq 1). \quad (18)$$

Substitution in (10) yields

$$3\sum_{j+k=n}(6k-2)(6k-5)\alpha_j b_{k-1}=2(n-1) \times\sum_{j+k=n+1}(6j-1)(6k-2)(6k-5)\alpha_j b_{k-1}, \quad (n\geqq 2). \quad (19)$$

The first equation of (18) gives the relation between α_1 and β_1. The remaining equations, when solved simultaneously in pairs, i.e., the second of (18) with the first of (19), etc., permit the other coefficients to be determined. The results are

$$\alpha_1=(3/5)(C\beta_1)^{\frac{1}{2}}, \quad \alpha_2/\alpha_1=-5/66, \quad \alpha_3/\alpha_1=131/30{,}600\cdots, \quad (20)$$

$$b_1=2/75, \quad b_2=29/12{,}305\cdots. \quad (21)$$

The numerical values of the first nine coefficients of each series are given in Table I.

It may be noted that V from Eq. (17) is zero for $\theta=0$, as it should. The constant β_1 may be expressed in terms of the anode voltage, since $V=V_a$ when $\theta=\alpha$. Hence

$$\beta_1=V_a/\tilde{f}(\alpha)=V_a/\alpha^{4/3}f(\alpha), \quad (22)$$

TABLE II. Numerical values of the various functions.[a]

θ deg	θ rad	f	$\tilde{f}$	F	T	T_1	h	h_1	ϕ	ϕ_1
0	0	1.00000	0	∞	0	0	1.0000	1.0000	0	0
5.73	0.1	1.00027	0.04643	99.9599	0.15014	0.10011	1.0030	1.0033	3°26′	3°49′
11.46	0.2	1.00107	0.11709	24.9600	0.30096	0.20088	1.0120	1.0134	6°52′	7°38′
17.19	0.3	1.00242	0.20132	11.0709	0.45328	0.30304	1.0275	1.0306	10°17′	11°26′
22.92	0.4	1.00433	0.29600	6.20964	0.60786	0.40734	1.0495	1.0554	13°41′	15°14′
28.65	0.5	1.00682	0.39956	3.95946	0.76555	0.51456	1.0788	1.0884	17°4′	19°0′
34.38	0.6	1.00991	0.51108	2.73700	0.92731	0.62569	1.1160	1.1307	20°24′	22°44′
40.11	0.7	1.01366	0.63002	1.99970	1.09422	0.74185	1.1622	1.1837	23°43′	26°27′
45.84	0.8	1.01809	0.75609	1.52105	1.26751	0.86439	1.2186	1.2492	26°58′	30°7′
51.57	0.9	1.02326	0.88915	1.19272	1.44868	0.99495	1.2868	1.3296	30°13′	33°45′
57.30	1.0	1.02925	1.02925	0.95768	1.63952	1.1356	1.3690	1.4281	33°18′	37°20′
63.03	1.1	1.03612	1.17653	0.78361	1.84211	1.2891	1.4669	1.5490	36°22′	40°50′
68.76	1.2	1.04399	1.33129	0.65102	2.05913	1.4589	1.5865	1.6981	39°21′	44°17′
74.49	1.3	1.05297	1.49396	0.54764	2.29384	1.6496	1.7305	1.8832	42°14′	47°38′
80.22	1.4	1.06320	1.66514	0.46539	2.55039	1.8677	1.9025	2.1154	45°0′	50°54′
85.95	1.5	1.07486	1.84562	0.39883	2.8339	2.1219	2.1127	2.4099	47°38′	54°3′
90.00	$\pi/2$	1.08410	1.97953	0.35905	3.0546	2.3304	2.2890	2.6679	49°25′	56°13′
91.68	1.6	1.08817	2.03637	0.34413	3.1512	2.4248	2.3694	2.789	50°8′	57°5′
97.42	1.7	1.10338	2.23867	0.29855	3.5106	2.7945	2.6848	3.284	52°28′	59°59′
103.14	1.8	1.12079	2.45407	0.26012	3.923	3.258	3.074	3.945	54°38′	62°43′
108.87	1.9	1.14081	2.68463	0.22732	4.404	3.854	3.559	4.84	56°36′	65°16′
114.60	2.0	1.1639	2.9328	0.19910	4.970	4.656	4.164	6.10	58°30′	67°38′
120.33	2.1	1.1906	3.2017	0.17455	5.61	5.76	4.92	7.90	59°50′	69°54′
126.06	2.2	1.2215	3.4951	0.15304	6.42	7.3(6)	5.87	10.(6)	61°5′	71°54′
131.79	2.3	1.258	3.818	0.1340	7.42	9.7(5)	7.0(0)	14.(6)	62°4′	73°40′
137.52	2.4	1.299	4.176	0.1172	8.6(4)	13.(6)	8.4(6)	20.(9)	62°49′	75°13′
143.25	2.5	1.348	4.575	0.1022	10.(1)	19.(7)	10.(4)	31.(1)	63°23′	76°32′
148.98	2.6	1.405	5.025	0.0888	12.(0)	(30)	(12.8)	(48)	63°59′	77°35′
154.71	2.7	1.47	5.53	0.0768	14.(3)	(49)	(15.7)	(77)		78°22′
160.44	2.8	1.55	6.11	0.0662	(17.1)	(84)	(19.7)	(127)		78°49′
166.17	2.9	1.64	6.77	0.0568	(20.4)	(147)	(25.5)	(210)		
171.90	3.0	1.74	7.51	0.0486	(24.2)	(253)	(35.5)	(343)		
177.63	3.1	1.85	8.34	0.0415	(28.4)	(396)	(57.5)	(528)		
180.00	π	1.89	8.71	0.0389	(30.2)	(480)	(75.5)	(611)		

[a] Values in parentheses are approximate.

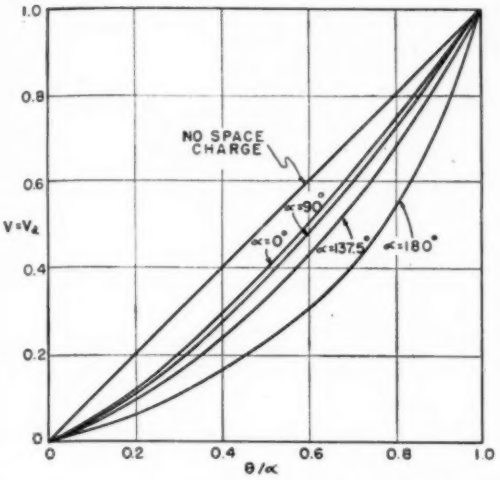


FIG. 2. Potential distribution between inclined plane electrodes (complete space-charge limitation). (The ordinate scale should be labeled V/V_a, not $V=V_a$.)

and

$$V/V_a=\bar{f}(\theta)/\bar{f}(\alpha). \tag{23}$$

The functions $f(\theta)$ and $\bar{f}(\theta)$ have been tabulated in Table II for values of θ between 0 and 180°. The effect of the angle between the electrodes on the potential variation is shown in Fig. 2, where V/V_a is plotted as a function of θ/α. Here the curve for $\alpha=0°$ is the same as that for plane electrodes, since $f(0)$ is unity. It is seen that as α is increased the potential is depressed. The effect is not very pronounced for angles less than 90°, but becomes progressively greater for larger angles.

Turning now to the magnification factor $h(\theta)$ given by Eq. (12), it is seen that this function could be determined by calculating g and its derivative g' from the α-coefficients and then performing a numerical integration. A simpler solution in series form can be obtained however, as shown by Walker, as follows. Let

$$h(\theta)=\exp\left(\int_0^\theta \frac{g}{g'}d\theta\right)=\exp\left(\sum_{n=1}^{\infty} a_n\theta^{2n}\right). \tag{24}$$

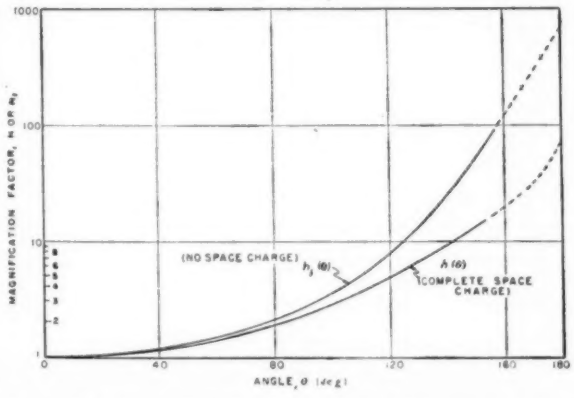


FIG. 3. Trajectory magnification factors for currents between inclined plane electrodes.

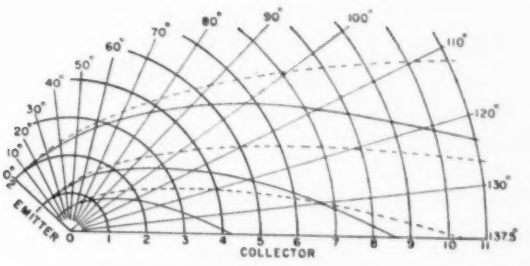


FIG. 4. Trajectories between inclined plane electrodes. Dashed line, absence of space charge; solid line, complete space charge.

Then by differentiation,

$$g=g'\sum_{n=1}^{\infty} 2na_n\theta^{2n-1}, \tag{25}$$

so that the coefficients a can be expressed in terms of the α-coefficients calculated earlier. The result is

$$3\alpha_n=2\sum_{j+k=n+1}(6j-1)k\alpha_j a_k, \quad (n\geqq 1). \tag{26}$$

Substitution yields

$$a_1=3/10, \quad a_2=3/220, \quad \cdots. \tag{27}$$

Walker has given the first two members of this series. Numerical values of the first nine coefficients are given in Table I. Values of $h(\theta)$ calculated from Eq. (24) are tabulated in Table II and plotted in Fig. 3. It is seen that the magnification increases very rapidly with increasing θ. For values of θ greater than about 150°, the values given are only approximate because more coefficients are needed to give adequate convergence of the series. Some examples of the trajectories have been plotted in Fig. 4, where for convenience θ has been limited to 137.5°.

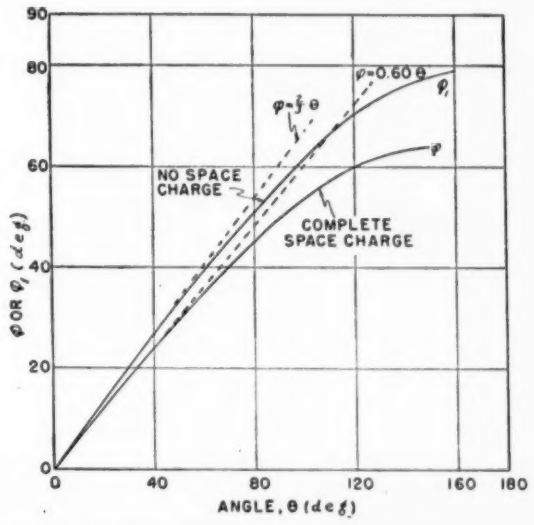


FIG. 5. Angle between trajectories and normals to equipotentials for inclined plane electrodes.

The angle of intersection, ϕ, of the trajectories with the equipotentials, given by Eq. (13), can also be conveniently calculated in terms of the a coefficients by use of Eq. (25). The results have been tabulated in Table II and plotted in Fig. 5. For small values of θ, $\varphi=0.600\theta$.

Let us now consider the current density corresponding to complete space-charge limitation. From Eq. (15) it is seen that the value of V'' is needed. By differentiation of (17) this may be written

$$V''=-\frac{4}{9}\frac{\beta_1}{\theta^{\frac{1}{3}}}f_2(\theta), \tag{28}$$

where

$$f_2(\theta)=1+\frac{35}{2}b_1\theta^2+52b_2\theta^4+\frac{209}{2}b_3\theta^6+\cdots. \tag{29}$$

Substituting Eqs. (17), (22), and (28) in (15) one obtains

$$J=\frac{1}{9\pi}\left(\frac{2e}{m}\right)^{\frac{1}{2}}\frac{V_a^{\frac{1}{2}}}{(r\alpha)^2}\frac{[f(\theta)]^{\frac{1}{3}}f_2(\theta)}{[f(\alpha)]^{\frac{1}{3}}}\cos\varphi, \tag{30}$$

where r is related to r_0 by Eq. (12). At the emitter ($\theta=0$) this expression reduces to

$$J_0=\frac{1}{9\pi}\left(\frac{2e}{m}\right)^{\frac{1}{2}}\frac{V_a^{\frac{1}{2}}}{(r_0\alpha)^2}\frac{1}{[f(\alpha)]^{\frac{1}{3}}}$$

$$=2.330\frac{V_a^{\frac{1}{2}}}{r_0^2}F(\alpha)\ (\mu\text{a/cm}^2), \tag{31}$$

where

$$F(\alpha)=\alpha^{-2}[f(\alpha)]^{-\frac{1}{3}}, \tag{32}$$

and the constant has been evaluated for electronic flow. As α approaches zero, Eq. (31) approaches the usual formula for parallel plane electrodes,[1,2] since $f(\alpha)$ approaches unity and $(r_0\alpha)$ is essentially the separation of the electrodes. The perveance* function $F(\theta)$ is tabulated in Table II and plotted in Fig. 6. Since $f(\alpha)$ is close to unity for small angles, J_0 varies roughly inversely as α^2 for small angles but more rapidly for large angles. For values of $\theta>70°$, F decreases roughly exponentially with θ.

Calculation of the current density at the collector from Eq. (30) is not satisfactory since the expression (29) for $f_2(\theta)$ does not converge very rapidly. This difficulty may be obviated, however, by simply recalling that the current emitted from a given area of the emitter is distributed over a larger area of the collector according to the magnification principle, so that

$$J_a=\frac{J_0}{h(\alpha)}=2.330\frac{V_a^{\frac{1}{2}}}{r^2}h(\alpha)\cdot F(\alpha)\ (\mu\text{a/cm}^2). \tag{33}$$

The function $h(\theta)\cdot F(\theta)$ has also been plotted in Fig. 6. It is seen that this function has a minimum value at about 105°.

Combining Eqs. (12) and (14), one obtains

$$t=r_0\int_0^\theta\frac{1}{g'}\exp\left(\int_0^\theta\frac{g}{g'}d\theta\right)d\theta. \tag{34}$$

The transit time could be determined by double numerical integration as indicated by this expression, but a series expansion is preferable. By differentiating (34) with respect to θ, taking the logarithm, and differentiating again, the result is the differential equation

$$\frac{d}{d\theta}(g't')=gt'. \tag{35}$$

Now write t as

$$t=A\theta^{\frac{1}{2}}(1+\sum_{n=1}^{\infty}p_n\theta^{2n})=A\theta^{\frac{1}{2}}p(\theta), \tag{36}$$

where p_0 has been taken as unity. The remaining p coefficients can be found in terms of the α's by substituting Eqs. (16) and (36) in (35). The resulting con-

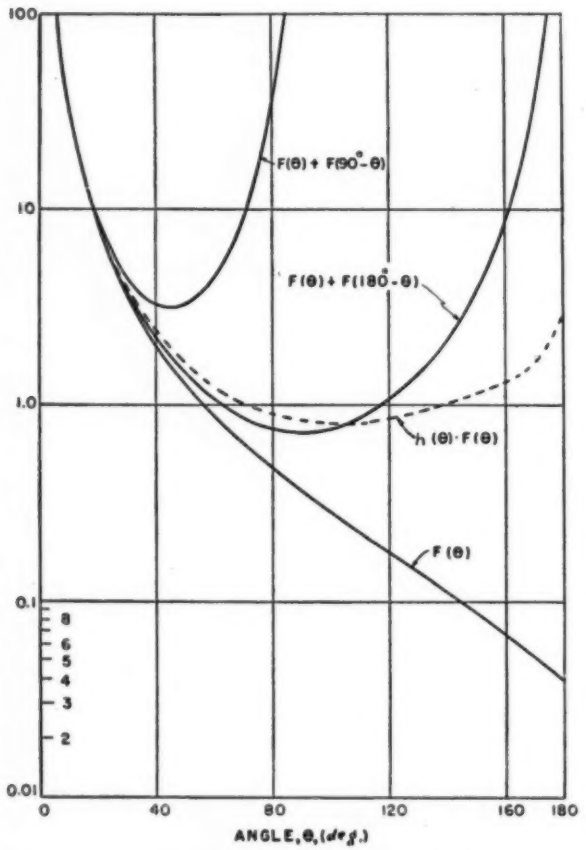


FIG. 6. Perveance functions for space-charge-limited currents between inclined plane electrodes.

* Strictly speaking, the perveance should include the $2.330/r_0^2$ factor also, since the perveance is, by definition, $J/V_a^{\frac{1}{2}}$.

nection formula is

$$3\sum_{j+k=n}(6k+1)\alpha_j p_k = 2n\sum_{j+k=n+1}(6j-1)(6k+1)\alpha_j p_k, \quad (p_0=1),(n\geqq 1). \quad (37)$$

The corresponding solutions are

$$p_1=1/15, \quad p_2=116/12{,}375\cdots. \quad (38)$$

The values of the first nine coefficients are given in Table I.

It will be seen that the above treatment has not yielded a value for the multiplier A in expression (36). This occurs because in the derivation of Eq. (35), differentiation of the logarithm causes the multiplicative constant to disappear. This constant may be evaluated, however, by the following reasoning. Let θ approach zero so that the series for g, g', and t may be broken off at the first term. Then the exponential term in (34) approaches unity and one obtains

$$t\to r_0\int_0^\theta \frac{d\theta}{(5/3)\alpha_1\theta^{\frac{3}{2}}}=\frac{9}{5}\frac{r_0}{\alpha_1}\theta^{\frac{1}{2}}. \quad (39)$$

The coefficient of $\theta^{\frac{1}{2}}$ here is the desired value of A. Substituting the value of α_1 from Eq. (20), and using Eqs. (7) and (22), one obtains

$$A=3r_0\left(\frac{m}{2e}\right)^{\frac{1}{2}}\frac{\alpha^{\frac{3}{2}}[f(\alpha)]^{\frac{1}{2}}}{V_a^{\frac{1}{2}}}. \quad (40)$$

It is convenient to introduce the quantity

$$t_0=\left(\frac{2m}{e}\right)^{\frac{1}{2}}=3.373\times 10^{-8}\ \text{sec volt}^{\frac{1}{2}}\ \text{cm}^{-1}, \quad (41)$$

where the numerical value applies to electronic flow. If now we let $\theta=\alpha$, since we are primarily concerned with the total transit time, Eq. (36) becomes

$$t_\alpha=\tfrac{3}{2}t_0\frac{r_0\alpha}{V_a^{\frac{1}{2}}}[f(\alpha)]^{\frac{1}{2}}p(\alpha)=\frac{r_0 t_0}{V_a^{\frac{1}{2}}}T(\alpha), \quad (42)$$

where

$$T(\alpha)=\tfrac{3}{2}\alpha[f(\alpha)]^{\frac{1}{2}}p(\alpha). \quad (43)$$

For small values of α it is seen that Eq. (42) reduces to the usual formula for parallel plane electrodes[8] if $r_0\alpha$ is interpreted as the electrode separation. Values of the function $T(\theta)$ are given in Table II and plotted in Fig. 7. The transit time increases more than linearly with the angle between the electrodes, as is to be expected because of the increased length of the trajectory due to the magnification factor h.

[8] K. R. Spangenberg, *Vacuum Tubes* (McGraw-Hill Book Company, Inc., New York, 1948), p. 195.

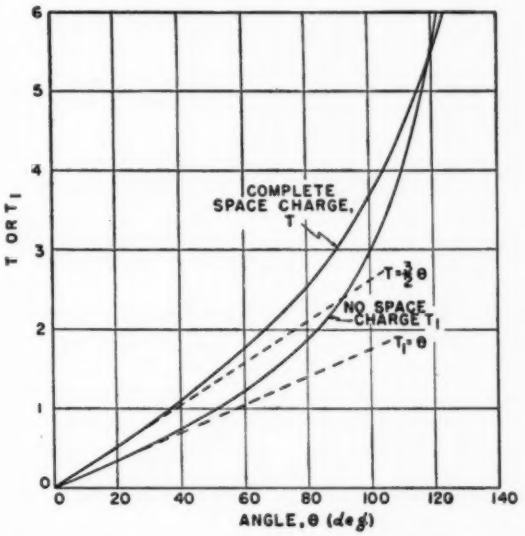


FIG. 7. Transit time functions for currents between inclined plane electrodes.

For values of θ less than α, Eq. (36) may be rearranged to give

$$\frac{t}{t_\alpha}=\left(\frac{\theta}{\alpha}\right)^{\frac{1}{2}}\cdot\frac{p(\theta)}{p(\alpha)}=\frac{T(\theta)}{T(\alpha)}\cdot\left[\frac{\bar{f}(\alpha)}{\bar{f}(\theta)}\right]^{\frac{1}{2}}. \quad (44)$$

4. SOLUTIONS FOR THE ABSENCE OF SPACE CHARGE

It is also of interest to consider current flow in the electrode arrangement considered in the absence of space charge. In this case the potential variation is obviously

$$V=V_a(\theta/\alpha). \quad (45)$$

The equation of continuity has no meaning here and only Eq. (6) remains to be satisfied. To prevent confusion, the symbol g used previously will be replaced by y for this case. Equation (6) becomes

$$y^2+y'^2=\left(\frac{2e}{m}\right)\left(\frac{V_a}{\alpha}\right)\theta=K\theta. \quad (46)$$

Following Walker, let

$$y=\sum_{n=1}^{\infty}c_n\theta^{(4n-1)/2}. \quad (47)$$

The connection formulas obtained upon substitution are

$$\left.\begin{aligned} 4K&=9c_1^2\\ 4\sum_{j+k=n}c_jc_k&=-\sum_{j+k=n+1}(4j-1)(4k-1)c_jc_k, \quad (n\geqq 2)\end{aligned}\right\}. \quad (48)$$

The corresponding solutions are

$$c_1=(2/3)K^{\frac{1}{2}}, \quad (c_2/c_1)=-2/21, \quad (c_3/c_1)=10/2079\cdots. \quad (49)$$

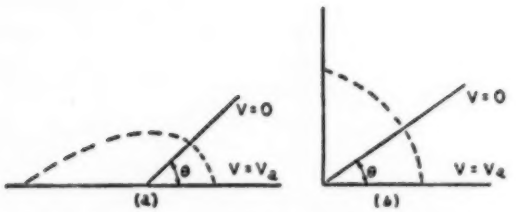


FIG. 8. Electrode arrangements for two-sided emitters.

Walker has given the first three members of this series. Values of the first eight coefficients are given in Table I.

Let the symbol h_1 designate the magnification function, given by Eq. (12), for this case. This function may be evaluated by a method similar to that used in the preceding section for the space-charge-limited case. Let

$$h_1(\theta)=\exp\left(\int_0^\theta \frac{y}{y'}d\theta\right)=\exp\left(\sum_{n=1}^{\infty} d_n\theta^{2n}\right), \quad (50)$$

so that

$$y=y'\sum_{n=1}^{\infty} 2nd_n\theta^{2n-1}. \quad (51)$$

Substitution yields

$$c_n=\sum_{j+k=n+1}(4j-1)kc_jd_k, \quad (n\geqq 1). \quad (52)$$

The solution of these equations yields

$$d_1=\tfrac{1}{3}, \quad d_2=4/189, \quad d_3=32/18{,}711, \quad \cdots. \quad (53)$$

Walker has given the first three members of this series. Values of the first eight coefficients are given in Table I.

The function $h_1(\theta)$ can now be calculated from Eq. (50). The results are tabulated in Table II and plotted in Fig. 3. For values of θ greater than about 150° the results are only approximate. A few trajectories are plotted in Fig. 4. It is seen that the magnification is always greater in the absence of space charge than in its presence and that $h_1(\theta)$ increases very rapidly as θ increases. The magnification is lower in the presence of space charge because the reduction of the potential by space charge causes the velocity to be everywhere decreased, and hence there is greater opportunity for the field to deflect the particles than in the absence of space charge. The angle ϕ_1 between the trajectories and the equipotentials can also be found from Eqs. (13) and (51) and the coefficients previously determined. The results are shown in Table II and Fig. 5. For small values of θ, $\phi_1=(2/3)\theta$.

Let us now consider the transit time in this case. The solution could be obtained by the method used in the space-charge-limited case, but since here V' is independent of θ, a simpler method, given by Walker, may be employed. Since V is a function of θ only, only the angular component of velocity is affected by the electric field. The field strength E is given in this case by

$$E=-(dV/rd\theta)=-(V_a/r\alpha). \quad (54)$$

The angular component of acceleration in polar coordinates (including the Coriolis acceleration) is

$$\frac{1}{r}\frac{d}{dt}(r^2\dot{\theta})=2\dot{r}\dot{\theta}+r\ddot{\theta}. \quad (55)$$

Substitution of (5b), (54), and (55) in Newton's second law of motion yields

$$\frac{d}{dt}(r^2\dot{\theta})=\frac{d}{dt}(ry')=\frac{e}{m}\frac{V_a}{\alpha}.$$

If $t=0$ is taken at $\theta=0$ then the transit time is

$$t_1=\frac{m}{e}\frac{\alpha}{V_a}ry'. \quad (56)$$

Making use of the coefficients of Eq. (49) and the value of K from (46), this expression becomes

$$t_1=\left(\frac{2m}{e}\cdot\frac{\alpha}{V_a}\right)^{\frac{1}{2}}r\theta^{\frac{1}{2}}\sum_{n=0}^{\infty}q_n\theta^{2n}=\left(\frac{2m}{e}\cdot\frac{\alpha}{V_a}\right)^{\frac{1}{2}}r\theta^{\frac{1}{2}}q(\theta), \quad (57)$$

where

$$q_n=\left(\frac{4n}{3}+1\right)\frac{c_{n+1}}{c_1}. \quad (58)$$

The values of the first eight coefficients are given in Table I.

If Eq. (41) is substituted in (56) and θ set equal to α, the resultant transit time may be written

$$t_{1\alpha}=t_0\frac{r\alpha}{V_a^{\frac{1}{2}}}q(\alpha)=\frac{r_0t_0}{V_a^{\frac{1}{2}}}T_1(\alpha), \quad (59)$$

where

$$T_1(\alpha)=\alpha h_1(\alpha)\cdot q(\alpha). \quad (60)$$

Values of the function T_1 have been tabulated in Table II and plotted in Fig. 7. For small values of α, the transit time in the space-charge-limited case is three-halves that of the space-charge-free case, as is the case for parallel planes,[8] but as α increases this ratio becomes smaller, because of the greater length of the trajectory in the absence of space charge. For $\alpha=118°$ the two transit times are equal.

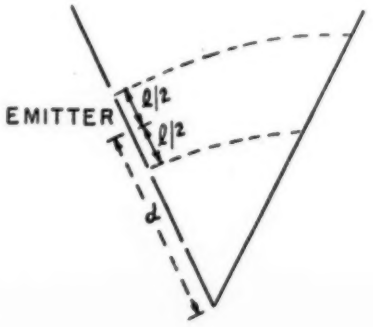


FIG. 9. Electrode arrangement for emitter of finite size.

For values of θ less than α, Eq. (57) may be rearranged to give

$$\frac{t_1}{t_{1\alpha}}=\left(\frac{\alpha}{\theta}\right)^{\frac{1}{2}}\frac{T_1(\theta)}{T_1(\alpha)}. \tag{61}$$

5. GEOMETRICAL CONSIDERATIONS

It is of interest to consider cases where both sides of the emitter are active. Two such cases are shown in Fig. 8(a) and 8(b) for a large plane collector and a right-angled collector, respectively. The functions giving the variation of perveance with angle in these cases are obviously $\{F(\theta)+F(180°-\theta)\}$ and $\{F(\theta)+F(90°-\theta)\}$, respectively. These functions are plotted in Fig. 6. It should be noted that the anode must be large enough to allow for trajectory magnification.

From Eq. (31) it is seen that the current density varies inversely as the square of the distance r_0 from the apex of the electrode arrangement, that is, the cathode loading is nonuniform. Consider now the total emission current density for a rectangular cathode of finite size as shown in Fig. 9. Guard plates are used to eliminate fringing fields† and it is also assumed that the anode is large enough to intercept all possible trajectories. Then for the contribution from an elemental area of the collector

$$dJ=2.330V_a^{\frac{1}{2}}F(\alpha)\frac{dr_0}{r_0^2},$$

so that

$$J=\int_{(d-l/2)}^{(d+l/2)}dJ=2.330V_a^{\frac{1}{2}}F(\alpha)\frac{G(l/d)}{d}\ (\mu\text{a/cm}), \tag{62}$$

where

$$G(l/d)=\frac{(l/d)}{1-(l/2d)^2}. \tag{63}$$

The function $G(l/d)$ has been plotted in Fig. 10. It is seen that for small values of l/d, G approaches l/d. To obtain the total current from Eq. (62) the value of J must obviously be multiplied by the length of the electrodes perpendicular to the plane of Fig. 9.

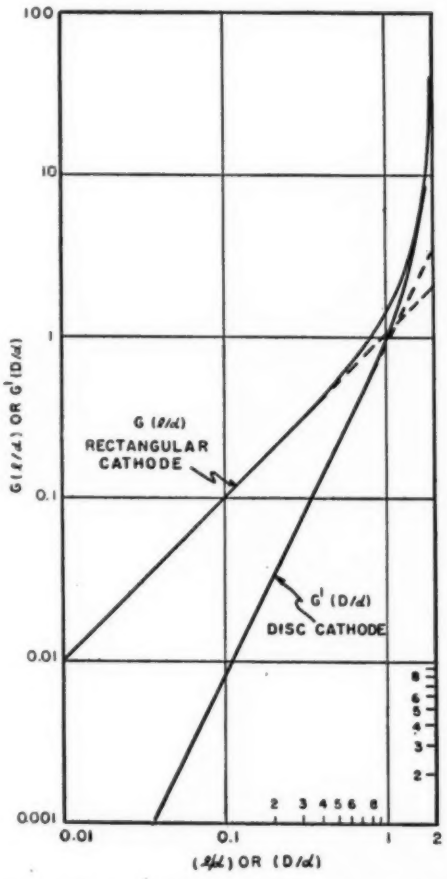


FIG. 10. Geometrical functions for rectangular and disk electrodes.

† Actually these "guard plates" should be shaped electrodes, as in a Pierce electron gun, to compensate for the missing regions of space charge, or should have the same emission properties as the cathode itself. This fact, pointed out to the writer by a referee for this paper, limits the application of the discussion in this section concerning finite emitters.

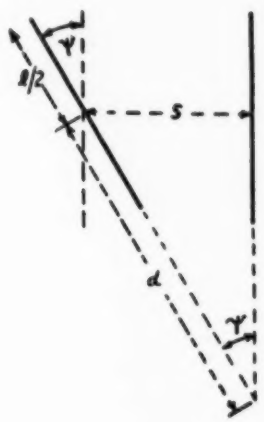


FIG. 11. Planar diode with tilted electrode.

Another interesting application is to consider the effect on the perveance of a planar diode, which presumably has parallel electrodes, brought about by accidental tilt of one of the electrodes. Consider the electrode arrangement shown in Fig. 11, where it is assumed that guard plates are actually present to prevent edge effects. One electrode is tilted about its center by an angle from the parallel condition. The situation is thus quite similar to that discussed above, and Eq. (62) is applicable if one substitutes

$$d=s/\sin\psi. \tag{64}$$

The current (per unit length) when $\psi=0$ may be written

$$J^0=2.330V_a^{\frac{1}{2}}l/s^2\ (\mu\text{a/cm}). \tag{65}$$

The ratio (J/J^0), which gives the effect of the tilt, thus becomes

$$\frac{J}{J^0}=\frac{F(\psi)}{(1/\sin\psi)^2-(l/2s)^2}. \tag{66}$$

Values calculated from this expression for small values of ψ and various values of l/s are given in Table III. It is seen that for small values of l/s the current ratio

TABLE III. Effect of electrode tilt on planar diode. (Rectangular electrode tilted about axis parallel to side.)

$\frac{l}{s}$	J/J^0 $\psi=5.73°$	$\psi=11.46°$	$\psi=17.19°$
0.1	0.9962	0.9852	0.9670
0.2	0.9963	0.9855	0.9677
0.5	0.9968	0.9876	0.9721
1.0	0.9987	0.9950	0.9884
1.2	0.9998	0.9993	0.9982
→ 1.24	1.0000	1.0003	1.0004
1.3	1.0003	1.0018	1.0038
2.0	1.0062	1.0256	1.0593
5.0	1.0624	1.308	2.129
10.0	1.327		

decreases as ψ increases, while for large values of l/s the reverse is true. For one particular value of l/s (roughly 1.24) the variation with ψ is essentially zero. It should be remembered that the rotation is assumed to occur about the center line of the cathode, i.e., the average value of the interelectrode spacing remains constant.

It is frequently of interest to construct a planar diode which will be relatively insensitive to changes in electrode alignment. The above considerations have shown that a rectangular electrode whose width is 1.24 times the interelectrode spacing is insensitive to tilt about a line passing through its center and parallel to its side. If it is desired that the diode characteristics should be independent of angular misalignment in an arbitrary direction, then the use of a disk-shaped emitter suggests itself. The situation in the plane of the emitter is then as shown in Fig. 12. Here

$$r_0=d-r\cos\gamma, \tag{67}$$

and the space-charge equation becomes

$$I=2.330V_a^{\frac{3}{2}}F(\alpha)G'(D/d)\ (\mu a), \tag{68}$$

where

$$G'(D/d)=\int_0^{D/2}\int_0^{2\pi}\frac{r\,dr\,d\gamma}{(d-r\cos\gamma)^2}. \tag{69}$$

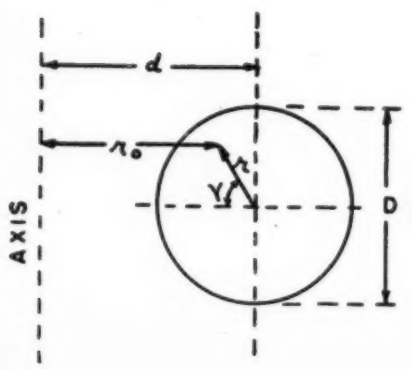


FIG. 12. Diode with disk emitter (in plane of emitter).

TABLE IV. Effect of electrode tilt on planar diode (disk cathode).

$\frac{D}{s}$	I/I^0 $\psi=5.73°$	$\psi=11.46°$	$\psi=17.19°$
0.1	0.9964	0.9852	0.9672
0.2	0.9956	0.9834	0.9636
0.5	0.9948	0.9868	0.9707
1.0	0.9980	0.9924	0.9840
1.4	0.9996	0.9997	0.9990
→ 1.42	0.9999	1.0001	0.9999
1.5	1.0005	1.0018	1.0039
2.0	1.0036	1.0152	1.0354
5.0	1.0452	1.215	1.714
10.0	1.232		

The integral is evaluated as follows:

$$\int_0^{2\pi}\frac{d\gamma}{(d-r\cos\gamma)^2}=2\pi d/(d^2-r^2)^{\frac{1}{2}},$$

$$\therefore\quad G'(D/d)=2\pi d\int_0^{D/2}\frac{r\,dr}{(d^2-r^2)^{\frac{3}{2}}}$$

$$=2\pi\left[\frac{1}{(1-(D/2d)^2)^{\frac{1}{2}}}-1\right]. \tag{70}$$

The function $G'(D/d)$ has also been plotted in Fig. 10. For small values of D/d, G' approaches $(\pi/4)(D/d)^2$.

Consider now the case when such a disk emitter is tilted relative to the collector so that Fig. 11 and Eq. (64) apply. The total current for $\psi=0$ in this case is

$$I^0=2.330V_a^{\frac{3}{2}}\pi D^2/4s^2\ (\mu a), \tag{71}$$

so that the effect of electrode tilt becomes

$$\frac{I}{I^0}=8F(\psi)\left(\frac{s}{D}\right)^2\left[\frac{1}{(1-(D\sin\psi/2s)^2)^{\frac{1}{2}}}-1\right]. \tag{72}$$

Values calculated from this expression for small values of ψ and various values of D/s are given in Table IV. It

TABLE V. Effect of electrode tilt on planar diode with disk cathode of optimum size.

ψ (radians)	I/I^0
0	1.0000
0.1	0.9999
0.2	1.0001
0.3	0.9999
0.4	0.9987
0.5	0.9970
0.6	0.9941
0.7	0.9871
0.8	0.9782
0.9	0.9615
1.0	0.9393
1.1	0.9059
1.2	0.8622
1.3	0.8065
1.4	0.7378
1.5	0.6590
$\pi/2$	0.5985

is seen that for $D/s=1.42$ the variation of current with ψ is essentially zero. This then is the criterion for a planar diode with disk emitter which shall be insensitive to angular misalignment. For this geometry Eq. (71) becomes

$$I^0=3.690V_a^{\frac{1}{2}}(\mu a). \tag{73}$$

It is also of interest to investigate the behavior of I/I^0 for this case for larger values of ψ. This has been done in Table V. It is seen that a tilt of 38° is required to change the current by one percent, while a tilt of 20° gives only one-tenth this effect.

6. CONCLUSION

It would seem worthwhile to check and extend the solutions given in this paper by means of a differential analyzer. Such an instrument, unfortunately, was not available for this purpose. A solution for the case of partial space-charge limitation (i.e., for an electric field at the emitter greater than zero) would be very desirable, but has not yet been found. This problem could be easily solved on an analyzer, however,

It would also be of interest to determine the variation of transit time with current for an electrode orientation of 118° where it is known that the transit time is the same for very small current flow and for complete space-charge limitation. This problem could also be studied with the differential analyzer.

The present method of attack does not allow consideration of initial velocities of emission. Some new method to take these into account would be of great value.

The numerical evaluations of the coefficients in the various series solutions employed have been checked by Mr. R. Gaines Young of this laboratory. His aid in this matter and also in critically reading the entire paper is gratefully acknowledged.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 2 FEBRUARY, 1952

Diffraction Measurements at 1.25 Centimeters*

RALPH D. KODIS,
Cruft Laboratory, Harvard University, Cambridge, Massachusetts
(Received September 10, 1951)

Measurements have been made, using an image plane technique, of microwave diffraction near obstacles whose characteristic dimension is comparable to a wavelength (1.2668 cm). Detector, obstacle, and source of radiation are mounted on one side of a horizontal 4-ft by 6-ft conducting plane, and the operator and all auxiliary apparatus are located on the other. Measured distributions of both the phase and amplitude of the electric field near long cylinders of brass and polystyrene were obtained, and for the conducting cylinders these are compared directly with the results of a suitably modified theory. The technique has been extended to diffraction by a semi-infinite conducting plane screen, and it was found in this case that considerable error is introduced by the point source. With the substitution of a line source the measurements are again in close agreement with the theory.

I. INTRODUCTION

WITH the development of microwave techniques many problems in electromagnetic diffraction have become of increasing interest. They deal with scattering by obstacles and apertures whose characteristic dimension is comparable to a wavelength, falling between the classical limits of optics on the one hand and the quasi-stationary state on the other. Only recently, however, have adequate experimental methods been developed for investigating these problems and for verifying the approximations that are usually involved in their analysis. As reported in the literature,[1,2] the techniques that have been employed involve arranging the optical components and the associated exciting and detecting apparatus in such a way that the effect of boundaries not included in the idealized theoretical problem is minimized or somehow accounted for. This procedure presents obvious difficulties since the extraneous boundaries whose effect is being taken into account are usually more complex than the one being investigated.

The set-up used in this experiment is designed to eliminate the possibility of interaction between the objects of interest and the apparatus required for diffraction measurements. To this end the components of the diffraction experiment are divided along a plane of symmetry and placed on one side of a horizontal conducting sheet; the observer and the measuring equipment are located in the field-free space on the other side. By thus minimizing the perturbations introduced into the primary diffracted field it becomes possible to make accurate and reliable measurements of its phase and amplitude.

* This research was made possible through support extended Cruft Laboratory, Harvard University, jointly by the Navy Department (ONR), the Signal Corps of the U. S. Army, and the U. S. Air Force, under ONR Contract N5ori-76, T. O. 1.

[1] C. L. Andrews, J. Appl. Phys. **22**, 465 (1951).

[2] Woonton, Carruthers, Elliott, and Rigby, J. Appl. Phys. **22**, 390 (1951).

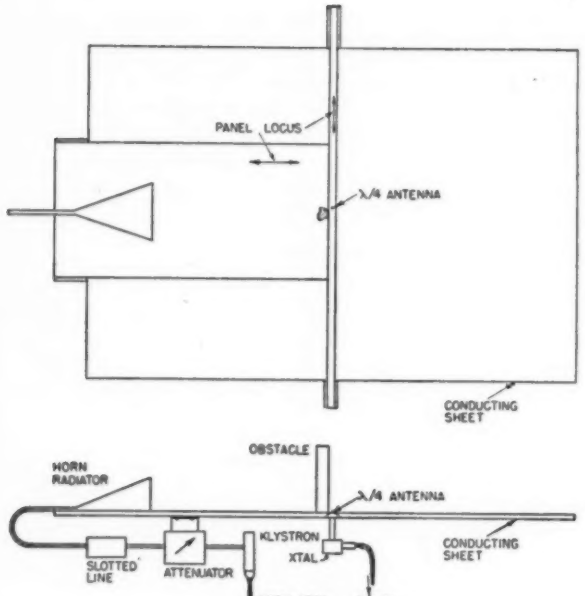


FIG. 1. Schematic diagram of the principal components used in the image-plane method of microwave diffraction measurements.

II. DESCRIPTION OF THE APPARATUS

In the experimental arrangement shown in Figs. 1 and 2 the image plane is approximated by a 4-ft by 6-ft sheet fabricated of duralumin panels. (At 24,000 megacycles the sheet is 96 wavelengths wide by 144 wavelengths long.) The elements of the scattering experiment are located on the top surface of the two movable panels in the sheet: the source of radiation and the obstacle are mounted on the broad sliding panel; a small detector is set in the narrow one. By suspending the exciting and detecting microwave circuits from their respective panels, two degrees of freedom are provided between obstacle and detector so that field distributions in the plane of the conducting sheet can be explored in a 5-in. by 10-in. (10-wavelength by 20-wavelength) rectangular area behind the obstacle.

The half-space above the image plane is maintained effectively free from obstructions by orienting the apparatus so that the incident radiation propagates toward an open window that is covered with a polyfoam screen (Fig. 2). Since the screen has a dielectric constant not much different from that of air and very low heat conductivity, it permits the scattered radiation to propagate out into free space without reflection while protecting the apparatus from the weather. Two squares of lossy material at the sides of the screen absorb some of the radiation incident upon them from the scattering obstacle and redirect the remainder out the window, thereby eliminating reflections from the window frame and adjacent wall areas. This arrangement, together with the relatively large sheet, reduces the residual standing waves in the area accessible to the probe to $\frac{1}{2}$ db or less.

The obstacle is illuminated by a horn radiator which flares in both E- and H-planes at an angle of 25 degrees to an aperture 16 wavelengths on a side. It is driven through a length of matched K-band wave guide by a $2K$-33 reflex klystron (free space wavelength $\lambda \sim 1.25$ centimeters or 0.5 inch) having a very closely regulated power supply modulated at 1000 cycles. Over the region accessible to measurement, the resulting incident field of the horn is almost exactly the same as that of a point source located at the apex of the horn 80 wavelengths from the obstacle (Fig. 3).

The detector consists of a quarter-wave antenna projecting through a .050-in. hole in the narrow transverse panel. This hole forms the outer conductor of a low impedance coaxial line which drives a wave guide terminated in a matched 1N26 crystal. The entire assembly is mounted on the undersurface of the panel to which a gear drive is attached. With an associated Ames gauge it is possible to locate the field detector to within $\pm .001$ in. over its entire range of travel.

Phase measurements are made in the conventional manner[3] by means of the interference minima produced by mixing the diffracted signal with a reference signal of variable phase. In this particular case the transmission line for the reference signal is made of wave guide sections connected by three rotating joints to provide flexibility without phase shift. Its length is adjusted so that the frequency dependence of the phase measurements is minimized. If, in addition, the relative amplitude of the two signals is properly adjusted,

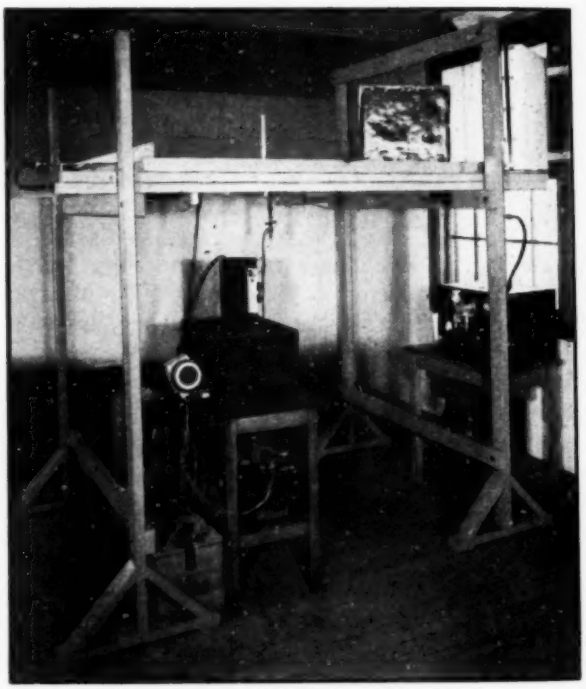


FIG. 2. General view of the microwave diffraction apparatus.

[3] C. G. Montgomery, *Techniques of Microwave Measurements* (McGraw-Hill, Book Company, Inc., New York, 1947) Radiation Laboratory Series, Vol. 11, p. 915.

phase measurements accurate to within 5 degrees are easily obtained.

III. COMPARISON OF MEASUREMENTS WITH THEORY

The extent to which the experimental arrangement described previously approaches the ideal conditions required by theory is best determined by direct comparison with a problem in electromagnetic scattering that can be solved exactly. Among the simplest of these is the diffraction of a plane wave by a conducting infinite cylinder of circular cross section. Since the results are well known in this case,[4] it is necessary here to repeat only what will be required later.

From the symmetry requirement imposed by the image plane we are led to choose the case of an incident plane wave with E-vector parallel to the cylinder axis. If this axis coincides with the z axis of a rectangular coordinate system whose xy plane is the plane of incidence, the entire field has only one component of the E-vector, and this component defines a scalar wave function that characterizes the problem completely. With harmonic time dependence $e^{-i\omega t}$, the wave function, $E_z(x, y)$, must satisfy the two-dimensional wave equation

$$(\nabla_{xy}^2+k^2)E_z(x, y)=0, \quad k=2\pi/\lambda, \tag{1}$$

at all points in space, subject to the boundary conditions that $E_z=E_z^{\text{inc}}+E_z^{\text{scatt}}$ vanish on the cylindrical surface and that the scattered wave satisfy the Sommerfeld radiation condition at infinity. Under these conditions it can be shown in a number of ways that if the incident plane wave is propagating in the positive x direction, as indicated in Fig. 4, then

$$E_z(x, y)=e^{ikx}$$

$$-\sum_{m=0}^{\infty}\epsilon_m i^{m+1}\sin\delta_m(ka)e^{-i\delta_m(ka)}H_m^{(1)}(kr)\cos m\theta, \tag{2}$$

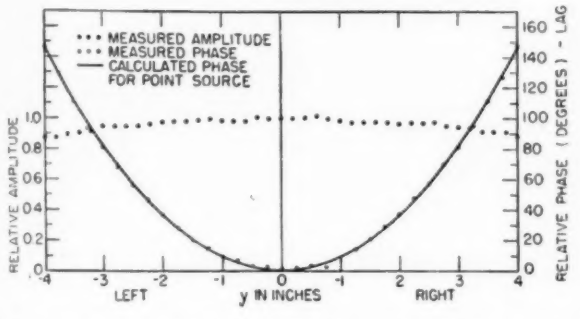


FIG. 3. Amplitude and phase of the incident radiation from the horn at 40 inches (80λ) from its apex.

[4] P. M. Morse, *Vibration and Sound* (McGraw-Hill Book Company, Inc., New York, 1936), p. 262 ff.

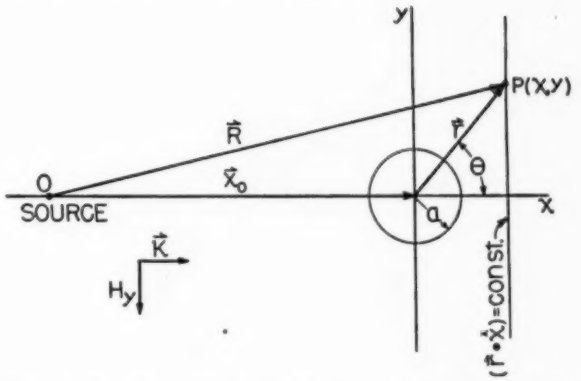


FIG. 4. Coordinates for cylindrical scattering. The line $(\hat{r}\cdot\hat{x})=$const is both the trace of the plane of uniform incident phase for an ideal plane wave and the locus of the measuring probe.

where the various quantities are defined as follows:

$$e^{ikx}=E_z^{\text{inc}},$$

$$\epsilon_m=\begin{cases}1 & \text{for } m=0\\ 2 & \text{for } m>0,\end{cases}$$

$$\tan\delta_m(ka)=-\frac{J_m(ka)}{N_m(ka)}$$

$$H_m^{(1)}(kr)=J_m(kr)+iN_m(kr)$$

$$r=(x^2+y^2)^{\frac{1}{2}}; \quad \tan\theta=y/x.$$

J_m and N_m are Bessel functions of the first and second kind, and $H_m^{(1)}$ is the Hankel function which is characteristic of outgoing cylindrical waves.

Before curves computed from (2) can be compared with corresponding measurements, it is necessary to take into account the fact that the radiating horn does not provide the incident plane wave assumed by the theory. To make the modification required by the actual point source, we notice first that the summation in (2) represents the scattered field which, in addition to satisfying the wave equation, must have the correct boundary values at $r=a$; that is

$$E_z^{\text{inc}}=-E_z^{\text{scatt}} \quad \text{at} \quad r=a. \tag{3}$$

This condition serves to determine the coefficients, $i^{m+1}\sin\delta_m(ka)\times e^{-i\delta_m(ka)}$, in Eq. (2); the remaining factors are required by (1) independently of the character of the incident field. Therefore in order to determine the form of the correct scattered wave we need only know the incident field at the cylindrical boundary. But it is readily shown that over this limited region (the largest value of $2a/\lambda$ is 3) the fields from a distant line source and from a plane wave are nearly identical, so the boundary values in (3) are not altered, and the *scattered* wave does not need to be corrected. We conclude that the form of (2) for the total diffracted field corresponding to a remote line source and an obstacle that is not too large (the exact condition turns

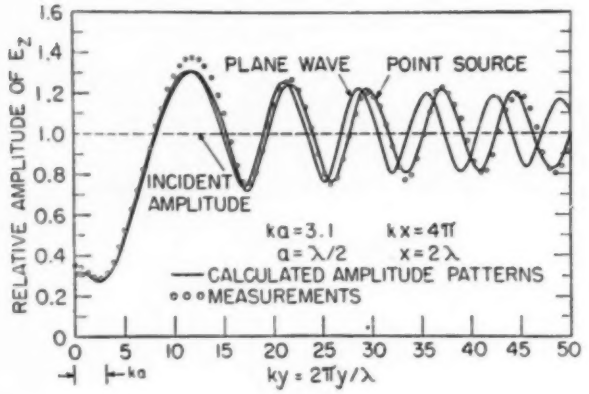


FIG. 5.

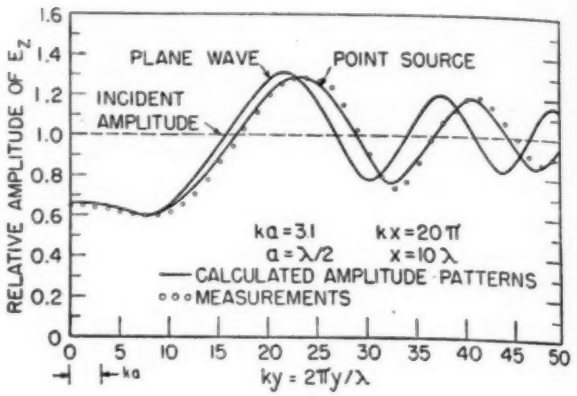


FIG. 7.

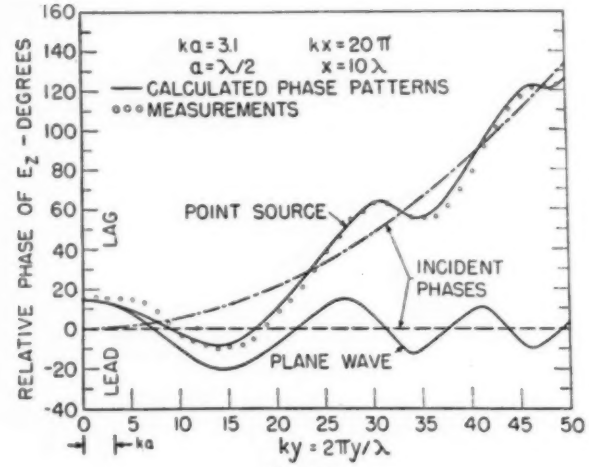


FIG. 6.

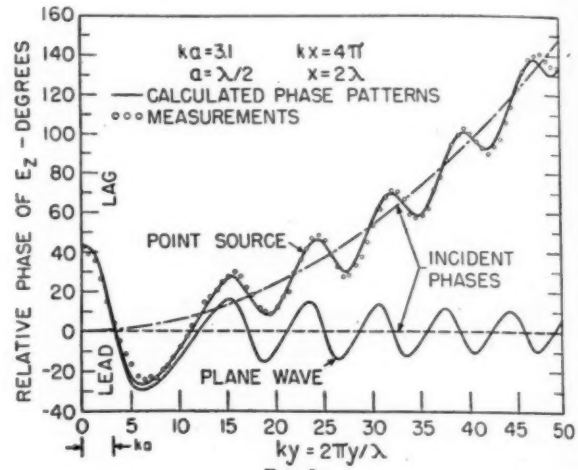


FIG. 8.

FIGS. 5–8. Comparison of calculated and measured diffraction patterns for a conducting cylinder one wavelength in diameter.

out to be $\frac{1}{2}a^2/x_0^2 \ll 1$) is

$$E_z(x, y) = e^{i\phi(x, y)} - \sum_{m=0}^{\infty} \epsilon_m i^{m+1} \sin\delta_m(ka) e^{-i\delta_m(ka)} H_m^{(1)}(kr) \cos m\theta, \quad (4)$$

where ϕ is the phase of the point source,

$$\phi(x, y) = k(x_0+x)\left[\left(1+\frac{y^2}{(x_0+x)^2}\right)^{\frac{1}{2}} - 1\right]. \quad (5)$$

Field distributions computed with Eq. (4) should agree with measurements, provided the fields produced by a point source and a line source do not differ significantly on the image plane. This approximation appears to be a reasonable one at distances not too far from the obstacle.

Representative results for brass cylinders 10 in. (20λ) long at two different distances are presented in Figs. 5 through 12. The quantities plotted are the amplitudes and phases of the total diffracted field normalized to the unperturbed incident values. For the amplitudes these are taken to be unity; for the phases both plane and cylindrical incident waves are assigned zero phase at the central point ($y=0$), and the phase difference between them at any other value of y is calculated from Eq. (5). Although the data shown are at the extremes of the possible parameter variations, agreement in both phase and amplitude is entirely satisfactory in every case.

IV. ADDITIONAL EXPERIMENTAL RESULTS

A. Dielectric Cylinders

When the conducting cylinder is replaced by one made of polystyrene, the resulting field distribution is much more complex, being characterized by rapid and large variations of both amplitude and phase. The theory required for a complete description of this field, while fundamentally the same as that of Sec. III, gives a correspondingly more complicated analytical result as a consequence of the boundary condition that must be satisfied on the surface of the dielectric. It is still true, however, that the boundary values are unaffected by the character of the incident wave as long as it is approximately plane in the vicinity of the obstacle. This being the case, if the cylindrical wave of Eq. (5) is subtracted from the measurements and a

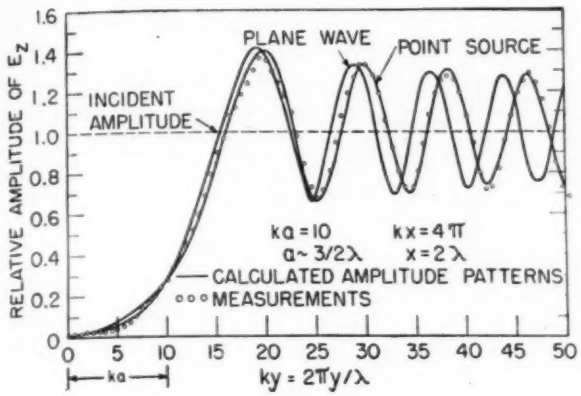


FIG. 9.

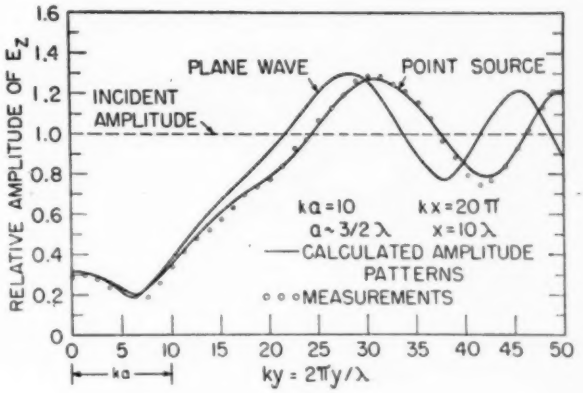


FIG. 11.

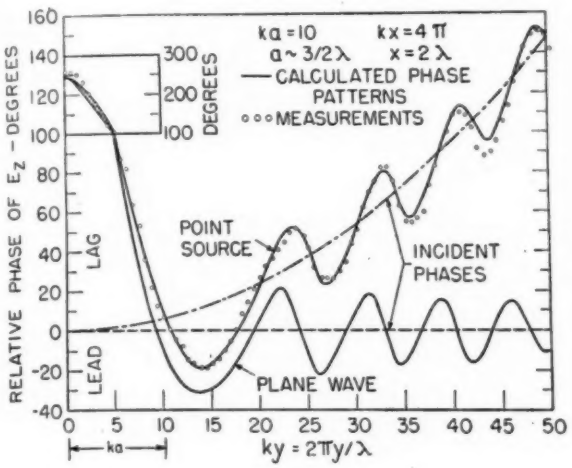


FIG. 10.

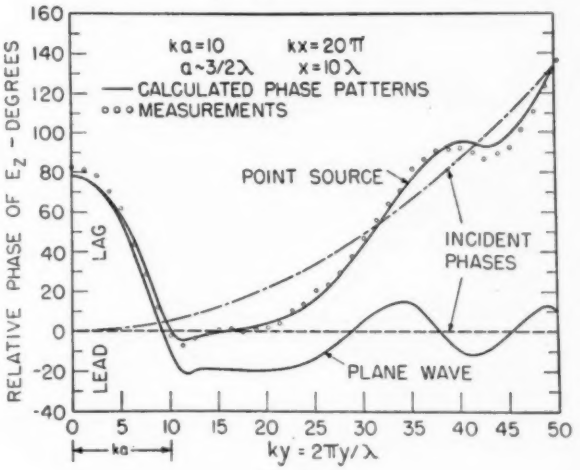


FIG. 12.

FIGS. 9–12. Comparison of calculated and measured diffraction patterns for a conducting cylinder three wavelengths in diameter.

plane wave added to the result, the transformed distributions will correspond to those obtained by computation from the plane wave theory.

A number of such distributions are shown in Figs. 13 through 16. Since the work involved in their computation would be prohibitive, they are not compared with the theory, but the curves illustrate clearly the regular, unambiguous results which are readily obtained with this apparatus.

B. The Diffracting Edge

It might appear from the foregoing results that if the approximation procedure involved in transforming distributions from one type of incident wave to another is to be valid, the diffracting obstacle must be small enough so that an incident cylindrical wave is very nearly plane over the entire obstacle cross section. Actually, the restriction is not so severe; the procedure may be applied to a larger obstacle when the principal contribution to its scattered field comes from the boundary values over a relatively small part of the cross section and when the obstacle can be located in the experimental setup so that the incident field is nearly plane over this limited region. The simplest example of a large diffracting boundary that meets these requirements is the semi-infinite conducting plane for which the significant boundary values are located near the diffracting edge.

In setting up this problem on the diffraction apparatus described previously the infinitely thin half-plane is approximated by a sheet of polished chrome-plated steel shim stock about 50 wavelengths square and 0.014 wavelength thick. The lower edge of the sheet is fastened securely in the image plane so that tension can be applied to the other edge, thus insuring that the plane is straight and vertical. In Fig. 4 the obstacle would be located over the negative y axis with the diffracting edge at the origin. The distributions are then measured along lines parallel to the screen.

To compare the results of these measurements with the Sommerfeld theory we make use of his formula[5] to get

$$E_z(r,\theta)=\frac{1}{\sqrt{2}}\exp[-i(\pi/4)]e^{ikr\cos\theta}\int_{-\infty}^{s}\exp[i(\pi t^2/2)]dt, \tag{6}$$

[5] M. Born, *Optik* (Verlag. Julius Springer, Berlin, 1933), p. 212.

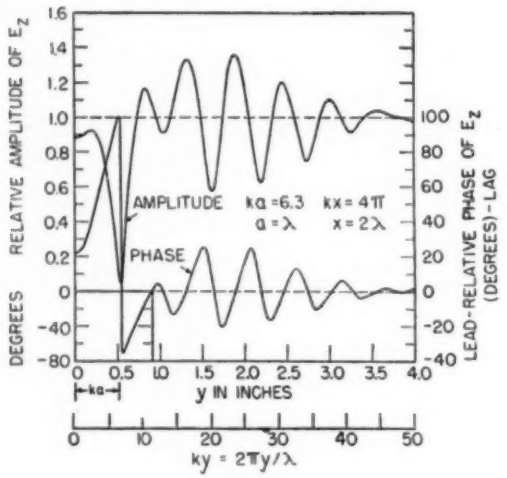


FIG. 13.

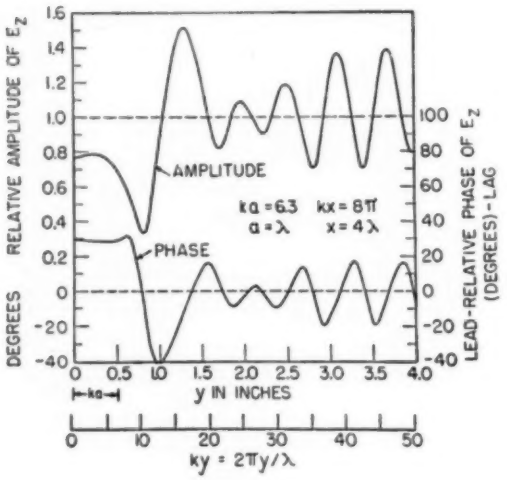


FIG. 14.

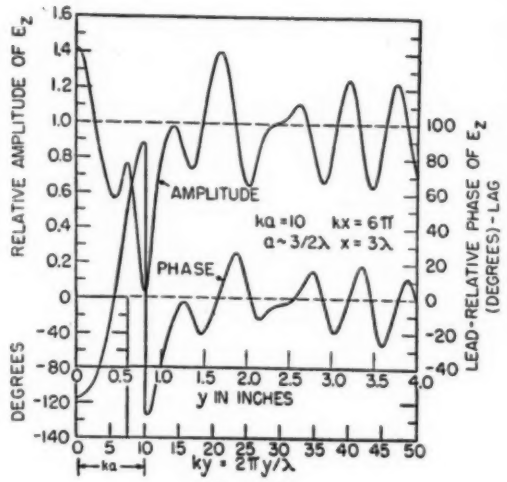


FIG. 15.

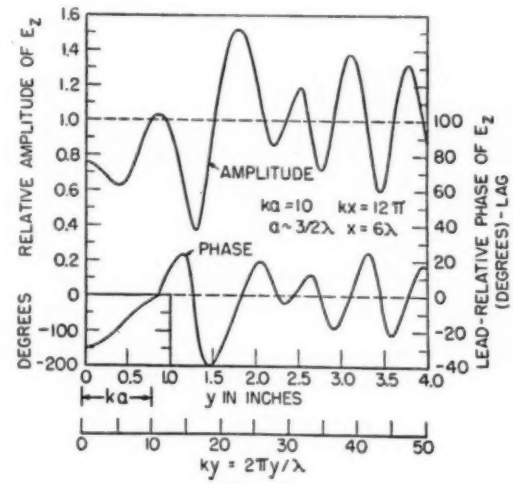


FIG. 16.

FIGS. 13–16. Measured diffraction patterns for polystyrene cylinders two and three wavelengths in diameter.

where

$$s=2\sqrt{\frac{kr}{\pi}}\cos\left(\frac{\theta-\pi}{2}\right).$$

$E_z(r, \theta)$ is the diffracted E vector for an incident wave polarized parallel to the edge and propagating in the direction of the normal to the screen. This field is evidently related to the Cornu spiral,

$$F(s)=\int_0^s \exp[i(\pi t^2/2)]dt;$$

in fact, along the probe locus where $r\cos\theta$ is constant, it is proportional to $[F(s)-F(-\infty)]$. It is therefore a relatively simple matter to obtain curves from (6) and, by means of the approximation procedure already discussed, to transform them into field distributions which are directly comparable to the measurements after these have been properly normalized to the incident wave. A typical pair of amplitude curves is shown in Fig. 17 which is characteristic of the best results obtainable with the single image plane. It will be noted that although the positions of the extreme values of the distribution are in accurate agreement, their magnitudes are considerably different. The discrepancy can be attributed to only two possible effects, since all other sources of error have been carefully eliminated: 1. The interaction between horn and screen. 2. The nonuniform illumination of the edge.

The second of these possibilities has been investigated with a recently constructed diffraction apparatus similar to the one already described except that it has two image planes which form a parallel plane transmission line supporting only the TEM mode at a wavelength of 3 cm. With the elements of the scattering problem sandwiched between these planes the equipment is especially well suited to the measurement of diffraction by obstacles which are either uniform or periodic in one direction. In particular, the flat feeding horn is imaged into a very close approximation of a uniform line source. So by inserting between the planes a vertical strip of lead foil, which is supported by a

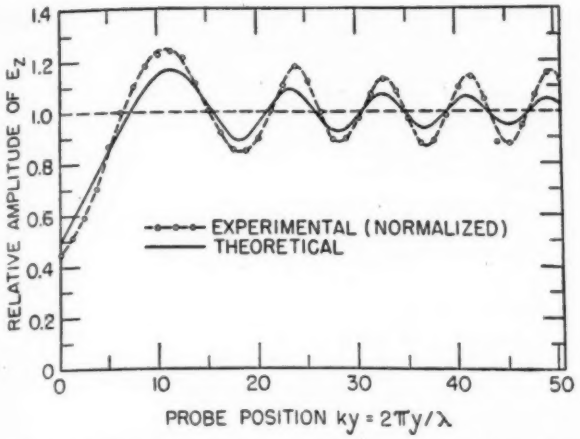


FIG. 17. Diffraction by a semi-infinite conducting plane. Comparison between measurements made on the *single* image plane and the theory corrected for a line source.

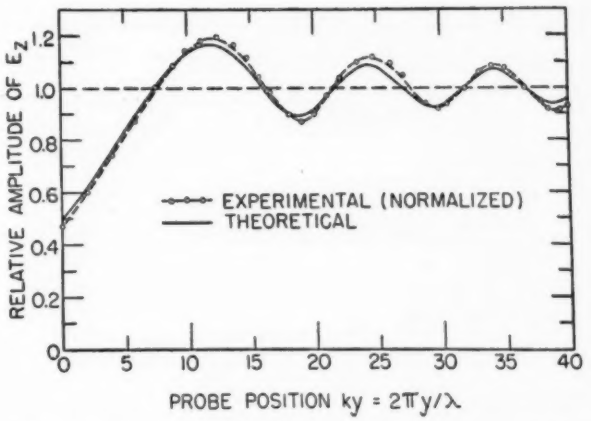


FIG. 18. Diffraction by a semi-infinite conducting plane. Comparison between measurements on the *double* image plane and the theory corrected for a line source.

polyfoam rod since it is only 0.005 in. (0.005λ) thick, the set-up comes very close to the idealized situation in the problem of the diffraction of a cylindrical wave by a semi-infinite conducting plane. Typical results obtained with this arrangement are given in Fig. 18. Almost complete agreement has been achieved, the small residual error probably being the result of the somewhat larger interaction on the double plane. (In terms of wavelengths the horn to plane separation is only one-half as great as that on the 1.25-cm equipment.)

V. CONCLUSION

We have seen that by employing an image plane technique it is possible to measure precisely the distribution in phase and amplitude of the *E*-field near a diffracting obstacle. If a two-dimensional distribution or the diffraction by a periodic array of infinite extent is being studied, the double plane gives the better results. However, if the obstacle (or aperture) lies within a region of the order of a few wavelengths in diameter, the single image plane must be used and in such cases gives results with the same degree of exactness. Three limitations are inherent in the method: 1. Arbitrary changes in the polarization and the direction of propagation of the incident radiation are not possible. 2. The obstacles must have a plane of symmetry. 3. Measurements are confined to the symmetry plane.

ACKNOWLEDGMENT

It is a pleasure to acknowledge the advice and encouragement given by Professor R. W. P. King who suggested this method of making diffraction measurements. The author is also grateful to Mr. C. Huang for carrying out the computations and the measurements for the diffraction by a plane screen and to Mr. R. V. Row who kindly repeated those measurements on his double image plane.

The Emission of Radiation from Diatomic Gases. III. Numerical Emissivity Calculations for Carbon Monoxide for Low Optical Densities at 300°K and Atmospheric Pressure*

S. S. PENNER, M. H. OSTRANDER,† AND H. S. TSIEN
Guggenheim Jet Propulsion Center, California Institute of Technology, Pasadena, California
(Received September 27, 1951)

Numerical emissivity calculations at 300°K and atmospheric pressure for nonoverlapping rotational lines have been carried out for CO using a dispersion formula for the line-shape representation. Use of the best available experimental data on integrated absorption and rotational line-width leads to calculated emissivities which are in excellent agreement with extrapolated empirical data published by Hottel and Ullrich. In particular, the theoretical dependence of emissivity on optical density, for small optical densities at 300°K, has been shown to follow experimental observations with satisfactory precision.

For small optical densities the calculated emissivity is found to be proportional to the square root of the assumed rotational line-width, thus emphasizing the need for accurate line-width determinations at elevated temperatures. The limits of validity of the treatment utilizing nonoverlapping rotational lines are defined by examining overlapping between adjacent weak and strong rotational lines.

The calculation of emissivities can be simplified by the use of approximate treatments. Thus absolute values of the emissivity can be predicted within 10 percent by utilizing a treatment for nonoverlapping, equally spaced, and equally intense lines, together with empirically determined values for the equivalent mean integrated absorption of the rotational lines of CO. A better analytic solution, which does not involve the assumptions of equal spacing and equal intensity of the rotational lines, has been obtained by utilizing asymptotic relations for large values of modified Bessel functions.

I. INTRODUCTION

THE general problem of emission and absorption of radiation involves two essentially distinct parts. One part of radiant heat transfer investigations is concerned with the determination of intensities if emissivities, absorptivities, and scattering coefficients are known. These studies are restricted to the problem of radiative transfer. For any defined geometric arrangement the radiative transfer problems can be formulated without difficulty. An exhaustive account of this work may be found in a recently published book by Chandrasekhar.[1] The other part of the general problem is the determination of emissivities, absorptivities, and scattering coefficients, which enter into the radiative transfer calculations. Until very recently, practical applications have been based almost entirely on the use of empirically determined emissivities. These emissivities are compiled, for instance, in McAdams' well-known treatise on heat transfer.[2] However, the problems involved in the theoretical calculation of emissivities were solved, in principle, a good many years ago. The results of these theoretical studies have been applied, to a limited extent, in the use of spectroscopic measurements for flame temperature determinations from line intensities[3] and for the analysis of absorption and emission of radiation in the atmosphere.[4] An important paper by Dennison,[5] written more than twenty years ago, contains many of the basic theoretical relations which have been used for numerical emissivity calculations on CO in the present studies.

In preceding publications an approximate method for carrying out emissivity calculations at elevated total pressures, leading to complete overlapping of rotational lines, has been described.[6–8] Limited comparison of calculated emissivities with experimental measurements has shown that the approximate method is, as expected, not accurately valid at atmospheric pressure for a diatomic emitter such as CO. In particular, the calculated dependence of emissivity on optical density for nonoverlapping rotational lines is not predicted correctly by a method based on the use of average absorption coefficients for entire vibration-rotation bands.[7] It is the purpose of the present investigation to demonstrate that the observed discrepancies between the previously proposed theory and the experiments can be removed by a more satisfactory analysis of the emission of radiation from vibration-rotation bands with rotational lines which do not overlap appreciably.

In Section II we present a summary of basic relations which are used for calculations on diatomic molecules with nonoverlapping rotational lines. Numerical emissivity calculations for CO at 300°K and a comparison with extrapolated experimental data are presented in Section III. The limits of validity of a treatment for nonoverlapping rotational lines are discussed in Section IV. In Section V approximate methods for performing emissivity calculations are considered under the conditions which are of interest for the present calculations.

* Supported, in part, by the ONR under contract Nonr-220 (03), NR 015 210, with the California Institute of Technology.

† This article uses, in part, the results of a thesis submitted by M. H. Ostrander in partial fulfillment of requirements for the degree of Aeronautical Engineer, California Institute of Technology, June, 1951.

[1] S. Chandrasekhar, *Radiative Transfer* (Clarendon Press, Oxford, 1950).

[2] H. C. Hottel, "Radiant heat transfer," in McAdams' *Heat Transmission* (McGraw-Hill Book Company, Inc., New York, 1942).

[3] For a review, containing numerous references to the original literature, see Am. J. Phys. **17**, 422, 491 (1949).

[4] W. M. Elsasser, Harvard Meteorological Studies No. 6, Blue Hill Observatory, Milton, Massachusetts, 1942.

[5] D. M. Dennison, Phys. Rev. **31**, 503 (1928).

[6] S. S. Penner, J. Appl. Phys. **21**, 685 (1950); L. E. Benitez and S. S. Penner **21**, 907 (1950).

[7] S. S. Penner, J. Appl. Mech. **18**, 53 (1951).

[8] S. S. Penner and D. Weber, J. Appl. Phys. **22**, 1164 (1951).

II. SUMMARY OF THEORETICAL RELATIONS

For diatomic molecules which do not possess a Q branch but a rotational line shape which may be represented by a dispersion formula, the spectral absorptions coefficient P_ω at the wave number ω resulting from energy transitions corresponding to the fundamental vibration-rotation band is[5,9,10]

$$P_\omega=(b/\pi)\sum_j\{S_j[(\omega-\omega_j)^2+b^2]^{-1}+S_{j-1}[(\omega-\omega_{j-1})^2+b^2]^{-1}\}, \quad (1)$$

where‡

$$S_j=(N_T\pi\epsilon^2/3\mu cQ)[(\omega_j/\omega^*)j\times\{\exp[-E(0,j)/kT]\}]FG(\omega^*/\omega_e), \quad (2a)$$

$$S_{j-1}=(N_T\pi\epsilon^2/3\mu cQ)[(\omega_{j-1}/\omega^*)j\times\{\exp[-E(0,j-1)/kT]\}]F'G'](\omega^*/\omega_e). \quad (2b)$$

Here N_T = total number of molecules per unit volume per unit of total pressure; μ = reduced mass of the diatomic molecule under discussion; c = velocity of light; ϵ = rate of change of electric moment with internuclear distance; b = rotational half-width in cm^{-1}; ω_e = wave number corresponding to an infinitesimal oscillation at the equilibrium interatomic distance; ω^* = wave number of the (forbidden) transition $j=0\rightarrow j=0$ and $n=0\rightarrow n=1$, where j and n denote, respectively, rotational and vibrational quantum numbers; ω_j = wave number corresponding to the transitions $j\rightarrow j-1$, $n=0\rightarrow n=1$; ω_{j-1} = wave number corresponding to the transitions $j-1\rightarrow j$, $n=0\rightarrow n=1$; $E(n,j)$ = energy of the nth vibrational and jth rotational level;

$$F=1+8\gamma j[1+(5\gamma j/4)-(3\gamma/4)];$$
$$F'=1-8\gamma j[1-(5\gamma j/4)-(3\gamma/4)];$$

$\gamma=h/8\pi^2Ic\omega_e$ where I represents the equilibrium moment of inertia of the radiating molecule;

$$G=1-\exp(-hc\omega_j/kT);$$
$$G'=1-\exp(-hc\omega_{j-1}/kT;$$

Q = complete internal partition function referred to zero as reference state.

At elevated temperatures the actual value of P_ω must be obtained by adding a small number of series of the type given in Eq. (1) corresponding to vibrational transitions of the form $n\rightarrow n+1$. Furthermore, contributions from vibration-rotation bands for which $n\rightarrow n+2$, $n\rightarrow n+3$, etc., become important for large optical densities. Detailed applications are discussed in reference 9.

The total equilibrium intensity of radiation emitted per unit surface area per unit time into a solid angle of 2π steradians in the wave number interval between ω and $\omega+d\omega$ by heated gases distributed uniformly and of optical density X (X = product of partial pressure p of emitter and radiation path length L) is given by the well-known relation

$$R(\omega)=R^0(\omega)[1-\exp(-P_\omega X)]. \quad (3)$$

Here $R^0(\omega)$ is the intensity of radiation emitted by a blackbody in the wave number interval between ω and $\omega+d\omega$ into a solid angle of 2π steradians per unit surface area per unit time. It is given in analytic form by the Planck blackbody distribution function. The engineering emissivity E is defined by the relation

$$E=\int_0^\infty R(\omega)d\omega/\sigma T^4, \quad (4)$$

where σ is the Stefan-Boltzmann constant.

Subject to the approximations and assumptions involved in the use of Eq. (1), it is evident that the problem of emissivity calculations reduces to the evaluation of integrals of the type appearing in Eq. (4). This integration can be accomplished if the width of the rotational lines is so small that they may be treated as being completely separated. The results obtained for nonoverlapping rotational lines are useful at low pressures.[4,9,11] The integrals can also be evaluated, approximately, if the pressures are so high that very extensive overlapping of the rotational lines occurs. The results obtained in this case are useful for emissivity calculations at elevated pressures in rocket combustion chambers.[6-8] Emissivity calculations at intermediate pressures are more difficult to carry out.

We proceed by considering the problem of emissivity calculations for nonoverlapping rotational lines. In order to evaluate the infinite integral

$$\int_0^\infty R(\omega)d\omega=\int_0^\infty R^0(\omega)[1-\exp(-P_\omega X)]d\omega, \quad (5)$$

for nonoverlapping rotational lines, the approximation is first made that $R^0(\omega)$ may be treated as constant in the wave number interval for which $\exp(-P_\omega X)$ is sensibly different from unity. Replacing $R^0(\omega)$ by an appropriate average value introduces negligibly small errors into the emissivity calculations for nonoverlapping rotational lines at low pressures, since the intervals $\Delta\omega$ for which $P_\omega X$ is significantly different from zero are exceedingly small. For example, at atmospheric pressure and room temperature $\Delta\omega$ is of the order of 0.3 cm^{-1} for CO. The change of $R^0(\omega)$ in wave number intervals of this size is less than 1 percent.

At low pressures the rotational lines are so narrow that the only term contributing to $R^0(\omega)$, for example, in the vicinity of the rotational line centered at ω_j is the term $S_j[(\omega-\omega_j)^2+b^2]^{-1}$. Hence we can break up the integration interval in such a way that each subinterval

[9] S. S. Penner, J. Chem. Phys. **19**, 272, 1434 (1951).

[10] J. R. Oppenheimer, Proc. Cambridge Phil. Soc. **23**, 327 (1926).

‡ Equations (2a) and (2b) differ from the corresponding relations obtained from the treatment of B. L. Crawford, Jr., and H. L. Dinsmore (J. Chem. Phys. **18**, 983, 1682 (1950)) through the occurrence of the factors F and F', respectively. These factors are nearly equal to unity for the values of j which are important for the present studies.

[11] R. Ladenburg and F. Reiche, Ann. Physik **42**, 181 (1913).

TABLE I. Integrated absorption for the rotational lines of the fundamental of CO at 300°K.

Transition $j \to j-1$	$10^3 A_j$ (cm^{-2}-atmos^{-1})	S_j (cm^{-2}-atmos^{-1})	Transition $j-1 \to j$	$10^3 A_{j-1}$ (cm^{-2}-atmos^{-1})	S_{j-1} (cm^{-2}-atmos^{-1})
1→0	5.528	2.064	0→1	5.572	2.081
2→1	10.71	4.000	1→2	10.88	4.062
3→2	15.28	5.705	2→3	15.64	5.841
4→3	19.02	7.101	3→4	19.62	7.326
5→4	21.79	8.135	4→5	22.65	8.457
6→5	23.51	8.780	5→6	24.64	9.200
7→6	24.24	9.050	6→7	25.58	9.551
8→7	24.00	8.960	7→8	25.54	9.535
9→8	22.97	8.577	8→9	24.63	9.199
10→9	21.32	7.960	9→10	23.04	8.603
11→10	19.23	7.180	10→11	20.94	7.819
12→11	16.88	6.310	11→12	18.53	6.919
13→12	14.45	5.397	12→13	15.99	5.969
14→13	12.07	4.509	13→14	13.46	5.025
15→14	9.851	3.678	14→15	11.06	4.131
16→15	7.857	2.934	15→16	8.889	3.319
17→16	6.127	2.288	16→17	6.986	2.608
18→17	4.673	1.745	17→18	5.370	2.005
19→18	3.490	1.303	18→19	4.039	1.508
20→19	2.551	0.9525	19→20	2.975	1.111
22→21	1.281	0.4783	21→22	1.622	0.6057
24→23	0.5924	0.2212	23→24	0.773	0.2885
26→25	0.2529	0.0944	25→26	0.340	0.1269
28→27	0.0998	0.0373	27→28	0.138	0.0516
30→29	0.0364	0.0136	29→30	0.052	0.0195
35→34	0.0021	0.0078	34→35	0.001	0.0052
40→39	0.0001	0.0037	39→40	0.000	0.0000

is centered about one of the wave numbers ω_j, ω_{j-1}.[9] Thus the following close approximation is obtained

$$\int_0^\infty R(\omega)d\omega=\sum_{\Delta\omega i}\left\{R^0(\omega_j)\int_{\Delta\omega i}[1-\exp(-P_jX)]d\omega + R^0(\omega_{j-1})\int_{\Delta\omega i}[1-\exp(-P_{j-1}X)]d\omega\right\}, \quad (6)$$

where $R^0(\omega_j)$ and $R^0(\omega_{j-1})$ denote, respectively, the spectral intensities of a blackbody radiator evaluated at the line centers corresponding to the indicated transitions. Similarly, P_j and P_{j-1} are the characteristic values of the spectral absorption coefficients valid if only the indicated transitions occurred. The subintervals $\Delta\omega_i$ are chosen conveniently to extend from a wave number midway between two line centers to the adjacent wave numbers which are located similarly. For nonoverlapping lines the error involved in replacing

$$\int_{\Delta\omega i}[1-\exp(-P_jX)]d\omega \text{ by } \int_{-\infty}^{\infty}[1-\exp(-P_jX)]d\omega,$$

etc., is negligibly small. The magnitude of this error increases with increasing optical density X and can be estimated readily by numerical calculations. Equation (6) now becomes

$$\int_0^\infty R(\omega)d\omega=\sum_j\left\{R^0(\omega_j)\int_{-\infty}^{\infty}[1-\exp(-P_jX)]d\omega + R^0(\omega_{j-1})\int_{-\infty}^{\infty}[1-\exp(-P_{j-1}X)]d\omega\right\}. \quad (7)$$

For fixed values of b the infinite integrals appearing in Eq. (7) are readily evaluated with the result[4,9,11]

$$\int_0^\infty R(\omega)d\omega=2\pi b\sum_{j=1}^{\infty}\{R^0(\omega_j)x_j \times\exp(-x_j)[I_0(x_j)+I_1(x_j)]+R^0(\omega_{j-1})x_{j-1} \times\exp(-x_{j-1})[I_0(x_{j-1})+I_1(x_{j-1})]\}, \quad (8)$$

where

$$x_j=S_jX/2\pi b, \quad x_{j-1}=S_{j-1}X/2\pi b,$$

and I_0, I_1 denote modified Bessel functions of the first kind. Calculations of the emissivity E from Eqs. (4) and (8), which neglect contributions from upper harmonics, can now be carried out by utilizing intensity[12] and line-width data[13] obtained at room temperature. The contribution to the emissivity E from the first overtone is obtained by adding to Eq. (8) a series which is of the same form as Eq. (8), except that the integrated absorption S_j is replaced by S_j' and S_{j-1} by S_{j-1}' where the primed quantities identify appropriate values for the vibrational transition 0→2.

III. REPRESENTATIVE EMISSIVITY CALCULATIONS FOR CO AT 300°K

For the present calculations the spectroscopic constants of Sponer were used.[14] Since these data have been

TABLE II. Integrated absorption for the rotational lines of the first overtone of CO at 300°K.

Transition $j \to j-1$	$10^3 S_j$ (cm^{-2}-atmos^{-1})	Transition $j-1 \to j$	$10^3 S_{j-1}$ (cm^{-2}-atmos^{-1})
1→0	1.428	0→1	1.440
2→1	2.768	1→2	2.811
3→2	3.948	2→3	4.042
4→3	4.914	3→4	5.070
5→4	5.629	4→5	5.852
6→5	6.076	5→6	6.367
7→6	6.263	6→7	6.610
8→7	6.201	7→8	6.598
9→8	5.935	8→9	6.365
10→9	5.508	9→10	5.953
11→10	4.968	10→11	5.411
12→11	4.366	11→12	4.788
13→12	3.735	12→13	4.130
14→13	3.120	13→14	3.477
15→14	2.545	14→15	2.859
16→15	2.030	15→16	2.297
17→16	1.583	16→17	1.805
18→17	1.208	17→18	1.388
19→18	0.9017	18→19	1.0437
20→19	0.6591	19→20	0.7688
22→21	0.3310	21→22	0.4191
24→23	0.1531	23→24	0.1996
26→25	0.0653	25→26	0.0878
28→27	0.0258	27→28	0.0357
30→29	0.0094	29→30	0.0135
35→34	0.0054	34→35	0.0036
40→39	0.0028	39→40	0.0000

[12] S. S. Penner and D. Weber, J. Chem. Phys. **19**, 807, 817, 974 (1951).
[13] S. S. Penner and D. Weber, J. Chem. Phys. **19**, 1351, 1361 (1951).
[14] H. Sponer, *Molekülspektren* (Verlag Julius Springer, Berlin, 1935).

revised recently,[15,16] all of the wave numbers of line centers, as well as the values of blackbody radiation from the line centers, will be somewhat in error. However, the over-all effect on the calculated values of emissivity is certainly negligibly small (less than 0.1 percent).

Assuming the validity of the ideal gas law and using recently published integrated intensity data[12] for the fundamental vibration-rotation band of CO, it is found that

$$(\omega^*/\omega_e)(N_T\pi\epsilon^2/3\mu c^2 Q)=371.9\ \text{cm}^{-2}\ \text{atmos}^{-1}.$$

The terms appearing in the large square brackets of Eqs. (2a) and (2b) are designated as A_j and A_{j-1}, respectively. Numerical values of A_j, A_{j-1}, and of S_j, S_{j-1} at 300°K for the fundamental vibration-rotation band of CO are listed in Table I.

The integrated absorption for the rotational lines of the first overtone, S_j' and S_{j-1}', are obtained, in close approximation, from S_j and S_{j-1}, respectively, by multiplying by the ratio of integrated absorption of the first overtone to the integrated absorption for the fundamental. The experimentally observed ratio[12] is $1.64/237=6.92\times10^{-3}$. Using this ratio, the results summarized in Table II are obtained for S_j' and S_{j-1}' at 300°K.

By utilizing the intensity data listed in Tables I and II, the emissivity has been calculated[17] from Eqs. (4) and (8) as a function of optical density and rotational half-width. The results of these calculations are summarized in Table III and plotted in Fig. 1, where they are compared with extrapolated empirical data reported by Ullrich.[18]

TABLE III. Calculated and extrapolated emissivities for CO at 300°K as a function of optical density and rotational half-width.

$T=300$°K, $b=0.06$ cm^{-1}, $X=pL$ (cm-atmos)	E (Calculated)	E (Extrapolated from reference 18)
0.1	3.62×10^{-4}	4.1×10^{-4}
0.5	9.30×10^{-4}	1.05×10^{-3}
2.0	1.92×10^{-3}	2.2×10^{-3}
6.0	3.43×10^{-3}	3.35×10^{-3}
$T=300$°K, $b=0.07$ cm^{-1}, $X=pL$ (cm-atmos)		
0.1	3.81×10^{-4}	4.1×10^{-4}
0.5	9.90×10^{-4}	1.05×10^{-3}
2.0	2.03×10^{-3}	2.20×10^{-3}
6.0	3.64×10^{-3}	3.35×10^{-3}
30.0	8.24×10^{-3}	5.8×10^{-3}
70.0	1.26×10^{-2}	7.9×10^{-3}
$T=300$°K, $b=0.08$ cm^{-1}, $X=pL$ (cm-atmos)		
0.1	4.13×10^{-4}	4.1×10^{-4}
0.5	1.08×10^{-3}	1.05×10^{-3}
2.0	2.21×10^{-3}	2.2×10^{-3}
6.0	3.90×10^{-3}	3.35×10^{-3}
30.0	8.84×10^{-3}	5.9×10^{-3}
70.0	1.35×10^{-2}	7.9×10^{-3}

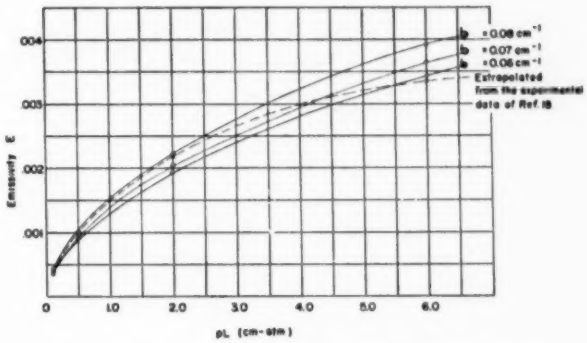


FIG. 1. Emissivity as a function of optical density at 300°K and atmospheric pressure.

Reference to the data listed in Table III and shown in Fig. 1 indicates that the calculated emissivity is proportional to the square root of the assumed value of the rotational half-width. Almost exact agreement between calculated and observed emissivities, for small optical densities, is obtained for $b=0.076$ cm^{-1} at atmospheric pressure. This value is in fair agreement with infrared line-width measurements on CO obtained by standard techniques.[13,19] We therefore arrive at the important conclusion that emissivities calculated from spectroscopic data are in quantitative agreement with empirically determined results, well within the claimed limits of accuracy of the latter measurements. The fact that the calculated emissivities appear to be too large for optical densities in excess of 2 cm-atmos is the result of failure of the assumption that no overlapping occurs between rotational lines. This matter will be discussed in Section IV.

Reference to the data listed in Table III and plotted in Fig. 1 shows that the emissivity E is very accurately a linear function of $(X)^{\frac{1}{2}}$. This result, together with the observation that E is proportional to the square root of the rotational half-width, is obviously in accord with Elsasser's well-known square-root law for nonoverlapping, equally intense, and equally spaced rotational lines.[4] Unfortunately, the simplified treatment[4] cannot be used to predict quantitatively (cf. Section VA for details) the slope of a plot of E *vs* $(X)^{\frac{1}{2}}$ in terms of the known basic spectroscopic constants.§

The results of the present calculations are of considerable practical importance. They clearly show that, for small optical densities, the emissivity is a sensitive function of rotational half-width. Thus we must conclude that not only is it not justified to extrapolate

[15] K. N. Rao, J. Chem. Phys. **18**, 213 (1950).
[16] Silverman, Plyler, and Benedict, Paper presented before the Symposium on Molecular Structure and Spectroscopy, Columbus, Ohio, 1951.
[17] Further details concerning this calculation may be found in the thesis of M. H. Ostrander.
[18] W. Ullrich, Dr. Sci. thesis, M.I.T., Cambridge, Massachusetts, 1935.
[19] L. A. Matheson, Phys. Rev. **40**, 813 (1932).
§ The same difficulty arises also if attempts are made to calculate the rotational half-width from experimental measurements of transmission. Compare reference 13.

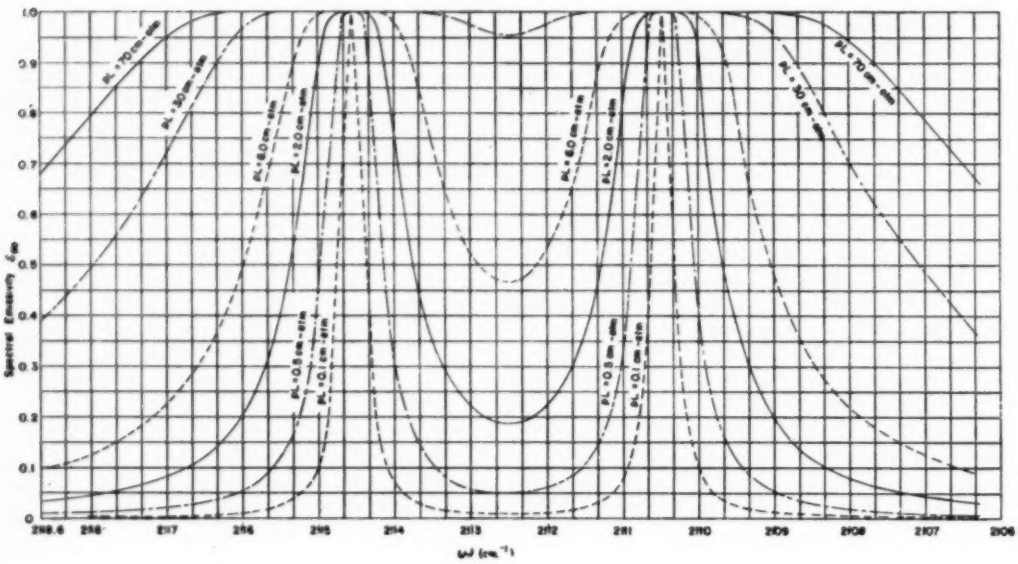


FIG. 2. Spectral emissivity ϵ_ω as a function of ω and optical density for two adjacent intense rotational lines ($n=0\rightarrow n=1$, $j=8\rightarrow j=7$ and $j=7\rightarrow j=6$) at 300°K and atmospheric pressure.

emissivities indiscriminately over temperature, pressure, or optical density intervals by using unsound simplified equations, but it is also essential to consider the problem with proper regard for the concentration of infrared-inert gases.

IV. LIMITS OF VALIDITY OF THE TREATMENT FOR NONOVERLAPPING ROTATIONAL LINES FOR CO[17]

To illustrate the effect of overlapping on P_ω, representative calculations involving the contributions to P_ω from two adjacent rotational lines have been carried out. In this case the correct expression for the spectral emissivity is evidently

$$\epsilon_\omega = 1-\exp[-(P_{\omega_1}+P_{\omega_2})]X, \qquad (9)$$

where P_{ω_1} and P_{ω_2} represent, respectively, the values of the spectral absorption coefficients arising from each of the two rotational lines separately. If the two rotational lines are separated sufficiently to permit neglect of overlapping, then the spectral emissivity is represented by the relation

$$\epsilon_\omega' = 1-\exp(-P_{\omega_1}X)+1-\exp(-P_{\omega_2}X). \qquad (10)$$

By evaluating the differences $\epsilon_\omega'-\epsilon_\omega$ from Eqs. (9) and (10) we can determine the error incurred in making numerical emissivity calculations, if overlapping between two adjacent rotational lines is ignored. The resulting expression,

$$\epsilon_\omega'-\epsilon_\omega = (P_{\omega_1}P_{\omega_2})X^2 - (P_{\omega_1}^2P_{\omega_2}+P_{\omega_1}P_{\omega_2}^2)X^3/2+\cdots, \qquad (11)$$

is useful for small values of $P_{\omega_1}X$ and $P_{\omega_2}X$.

Emissivity calculations for two rotational lines based on the correct expression given in Eq. (9) can be used as a qualitative guide in ascertaining the limits of validity of theoretical calculations based on a treatment for nonoverlapping rotational lines. Representative values of ϵ_ω calculated from Eq. (9) are shown in Figs. 2 and 3 as a function of ω with X treated as a variable parameter. Reference to Fig. 2 indicates that the emissivity at a point midway between the rotational line centers which are characterized by the transitions $n=0\rightarrow n=1$, $j=8\rightarrow j=7$, and $j=7\rightarrow j=6$, respectively, remains very small (i.e., $\epsilon_\omega<0.2$) for $X=2$ cm-atmos. Since the indicated transitions correspond to the two most intense rotational lines, it appears justifiable to conclude that emissivity calculations based on the use of Eq. (8) will yield valid results at least for $X\leq 2$ cm-atmos. Calculations similar to those shown in Fig. 2 for intense adjacent rotational lines are shown in Fig. 3 for weak rotational lines, namely, for the lines corresponding to the transitions $n=0\rightarrow n=1$, $j=19\rightarrow j=18$, and $j=18\rightarrow j=17$. As is to be expected, the emissivity ϵ_ω is represented more adequately by Eq. (10) for larger values of X than it was for the more intense rotational lines described in Fig. 2.

Although the precise evaluation of errors arising from the use of Eq. (8) is difficult to carry out, it is evident from the data described in Figs. 2 and 3 that the calculated values of the emissivity are reliable at least to $X=2$ cm-atmos and probably to somewhat larger values. This conclusion is borne out by the comparison between calculated and observed emissivities described in Section III.

V. APPROXIMATE EMISSIVITY CALCULATIONS FOR DIATOMIC MOLECULES WITH NONOVERLAPPING ROTATIONAL LINES

The numerical work involved in the emissivity calculations indicated in Section III is very heavy. It is thus worthwhile to consider the possibility of developing approximate analytical expressions which avoid much of the tedious computations. In this section we shall present two such approximations, in the order of increasing accuracy.

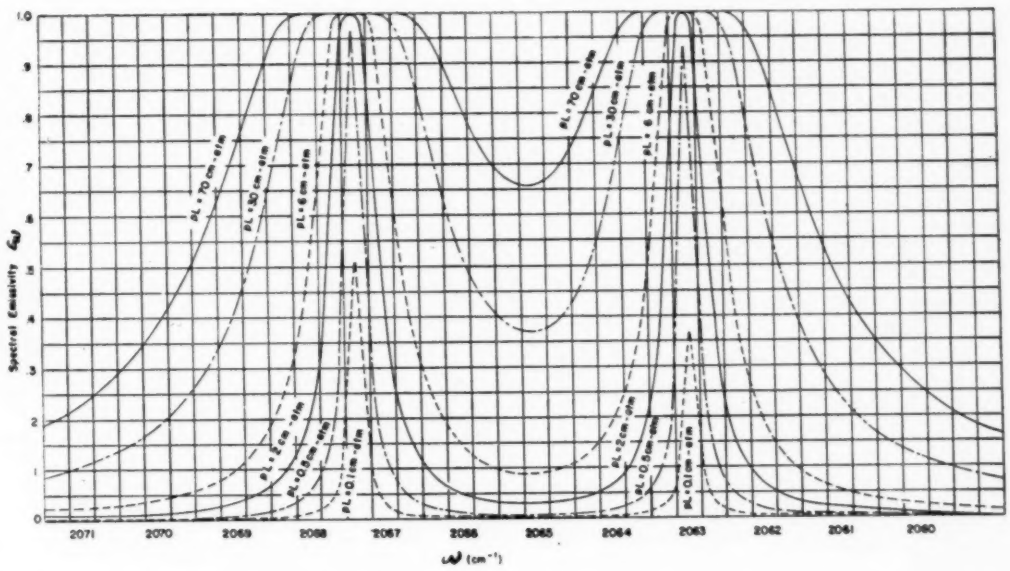


FIG. 3. Spectral emissivity ϵ_ω as a function of ω and optical density for two adjacent weak rotational lines ($n=0 \rightarrow n=1, j=19 \rightarrow j=18$ and $j=18 \rightarrow j=17$) at 300°K and atmospheric pressure.

VA. Emissivity Calculations for CO at 300°K for Nonoverlapping Rotational Lines, Assuming Equal Spacing and Intensity of Lines

From Eqs. (3) and (4) it is apparent that

$$E \simeq (\bar{R}^0/\sigma T^4)\int_0^\infty [1-\exp(-P_\omega X)]d\omega, \qquad (12)$$

where $\bar{R}^0$ represents an average value for $R^0(\omega)$ in the wave number interval for which P_ω is sensibly different from zero. The principal contribution to radiant heat transfer arises in the present problem from the fundamental vibration-rotation band. For moderate values of the optical density and nonoverlapping, equally spaced, and equally intense rotational lines it is well-known that[4]

$$\int_0^\infty [1-\exp(-P_\omega X)]d\omega = \Delta\omega_F \cdot 2(\bar{S}_F b X)^{\frac{1}{2}}/q. \qquad (13)$$

Here $\Delta\omega_F$ is the effective bandwidth,[6] $\bar{S}_F$ represents the integrated absorption of the equally intense rotational lines[4,13] for the fundamental vibration-rotation band of CO, and q is the spacing of the equally spaced rotational lines. Combining Eqs. (12) and (13) leads to the result

$$E = 2(\bar{R}^0 \Delta\omega_F/\sigma T^4)(\bar{S}_F b X)^{\frac{1}{2}}/q. \qquad (14)$$

As is known from the known limits of validity of the assumption that the lines are equally spaced and equally intense, Eq. (14) predicts accurately all of the qualitative features of an exact numerical solution for small values of X. Thus

$$\partial E/\partial(X)^{\frac{1}{2}} = \text{constant for fixed } b,$$

and

$$\partial E/\partial(b)^{\frac{1}{2}} = \text{constant for fixed } X.$$

Actually, we can do better than predict the functional form of E. Thus it has been shown[13] that for CO at 300°K

$$\bar{S}_F = 1.9 \text{ cm}^{-2} \text{ atmos}^{-1}. \qquad (15)$$

Also $\Delta\omega_F \simeq 250$ cm^{-1} extending from 2000 to 2250 cm^{-1}, $q=4$ cm^{-1}, and $b=0.08$ cm^{-1}. Hence, Eq. (12) becomes at 300°K

$$E \simeq 1.48\times10^{-3}(X)^{\frac{1}{2}}, \qquad (16a)$$

where X is in cm-atmos, $b=0.08$ cm.

The emissivity E calculated for CO at 300°K for a rotational half-width of 0.08 cm^{-1} (cf. Table III) is plotted as a function of $(X)^{\frac{1}{2}}$ in Fig. 4. Reference to Fig. 4 shows that

$$E = 1.64\times10^{-3}(X)^{\frac{1}{2}}$$
$$(b=0.08 \text{ cm}^{-1}, 0.1 \leq X \leq 2 \text{ cm-atmos}). \qquad (16b)$$

Comparison of Eqs. (16a) and (16b) shows that we can predict absolute emissivity values with an accuracy of about 10 percent without performing numerical calculations.

Considering that the objective of the present investigations is the theoretical calculation of emissivities from spectroscopic data, the following objections may be raised to the use of Eq. (14): (1) the absolute error involved in the use of Eq. (14) cannot be estimated without performing accurate numerical calculations; (2) Eq. (15) is an empirical relation whence it follows that the dependence of $\bar{S}_F$ on temperature and total pressure cannot be predicted accurately. Nevertheless, particularly in view of the fact that Elsasser[4] has obtained an integral representation for partial overlapping between equally spaced and equally intense rotational lines, the use of simplified treatments of the type given in Eq. (14) may be indicated in some cases. Thus we note that as the temperature is increased $\bar{P}_F$ must approach the ratio of integrated absorption to bandwidth,[8] a quantity which can be calculated[6] without difficulty.

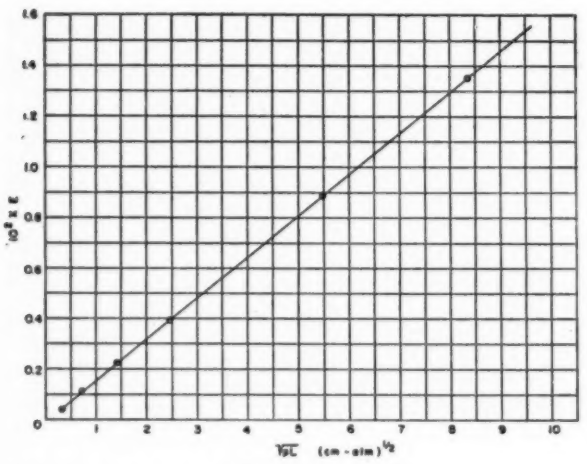


FIG. 4. Calculated emissivity E as a function of the square root of the optical density at 300°K and atmospheric pressure ($b=0.08$ cm^{-1}).

VB. Emissivity Calculations for CO at 300°K for Nonoverlapping Rotational Lines Using Asymptotic Expressions for the Modified Bessel Functions

The emission functions

$$f(x)=x\exp(-x)[I_0(x)+I_1(x)], \tag{17}$$

appearing in Eq. (8) have been studied and evaluated by Elsasser[4] who has emphasized the fact that, even for moderate numerical values of x, asymptotic expressions for the modified Bessel functions yield very good approximations to the value of $f(x)$. However, this result has not heretofore been utilized in connection with accurate treatments on nonoverlapping rotational lines in which proper allowance is made for the change in intensity from one rotational line to another. It is the purpose of the present discussion to evaluate the integral appearing in Eq. (8) and to determine the value of the emissivity E from Eq. (4) under conditions where the asymptotic expressions for the modified Bessel functions can be used.

Before describing the details of the integration it may be in order to emphasize the fact that the results of the present approximate treatment must be used with caution since the analysis rests on two seemingly contradictory assumptions. Thus overlapping between rotational lines is minimized by small values of x (compare Section IV), whereas the asymptotic expressions for the modified Bessel functions apply most accurately for large values of x. Fortunately a range of values of x exists (particularly at low pressures where the rotational half-width is small) for diatomic molecules with relatively distant rotational spacing, where both restrictive conditions are satisfied closely. In particular, for the problem under discussion, it has already been shown that overlapping between rotational lines is unimportant for optical densities smaller than about 2 cm-atmos (compare Sections III and IV). Furthermore, the range of values of integrated absorption for the rotational lines of the fundamental vibration-rotation band (cf. Table I) is such that the error made by setting[4]

$$f(x)=2(x/2\pi)^{\frac{1}{2}} \tag{18}$$

is small for most of the rotational lines, provided the optical density is greater than about 0.25 cm-atmos.‖ Hence, it is to be expected that a useful approximate solution will be obtained at least for 0.25 cm-atmos $\leq pL\leq 2.0$ cm-atmos. The approximate treatment utilizing asymptotic expressions for the modified Bessel functions is, of course, not applicable to calculations on the first overtone because of the small values of S_j', S_{j-1}' (cf. Table II). However, as was noted in Sections III and VA, the contribution to the emissivity E from the first overtone is negligibly small in the present problem.

Utilizing the approximation expressed by Eq. (18) in Eq. (8) it is found from Eq. (4) that

$$E=30(u/\pi)^4\sum_{j=1}^{\infty}\left\{\frac{(\omega_j/\omega^*)^3}{[\exp(u\omega_j/\omega^*)]-1}[bS_jX/(\omega^*)^2]^{\frac{1}{2}}+\frac{(\omega_j/\omega^*)^3}{[\exp(u\omega_{j-1}/\omega^*)]-1}[bS_{j-1}X/(\omega^*)^2]^{\frac{1}{2}}\right\}, \tag{19}$$

where explicit relations have been introduced for the Stefan-Boltzmann constant and, by using the well-known Planck distribution law, for the spectral radiation intensity emitted from a blackbody. Here we have adopted the notation of Mayer and Mayer[20] with

$$u=hc\omega^*/kT=h\nu^*/kT.$$

In terms of the customary notation for the spectroscopic constants[21] we now introduce the following close approximations, neglecting terms involving $(\gamma j)^3$ and higher,

$$\omega_j/\omega^*=1-2\gamma j-(\delta/\gamma)\gamma^2j^2, \tag{20}$$

$$F^{\frac{1}{2}}=1+4\gamma j-3\gamma^2j^2, \tag{21}$$

$$\begin{aligned}G^{\frac{1}{2}}[\exp(u\omega_j/\omega^*)-1]^{-1}&=e^{-u}(1-e^{-u})^{-\frac{1}{2}}\{1+\gamma ju[2+e^{-u}(1-e^{-u})^{-1}]\\&+\gamma^2j^2[u(2u+\delta/\gamma)+ue^{-u}(1-e^{-u})^{-1}(3u+\delta/2\gamma)\\&+(3/2)u^2e^{-2u}(1-e^{-u})^{-2}]\},\end{aligned} \tag{22}$$

‖ By utilizing the known values of $f(x)$ and of S_j, S_{j-1} it is readily shown that the error involved in using asymptotic forms for the Bessel functions is

less than 5 percent for $1\leq j\leq 18$ for $X=2$ cm-atmos,
less than 5 percent for $2\leq j\leq 16$ for $X=1$ cm-atmos,
less than 15 percent for $2\leq j\leq 14$ for $X=0.25$ cm-atmos, etc.

Since most of the contributions to E result from the rotational lines with $1\leq j\leq 16$, the preceding statement suggests the range of validity of the present treatment.

[20] J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley and Sons, Inc., New York, 1940).

[21] A compilation of numerical values may be found on p. 468 of reference 20.

$$\exp\{-[E(0,j)-E(0,0)]/2kT\}$$
$$=[\exp(-\gamma j^2 u/2)][1-(\gamma j u/2)+(\gamma^2 j^2 u^2/8)], \quad (23)$$

and[20]

$$Q\exp[E(0,0)/kT]$$
$$=(\gamma u)^{-1}(1-e^{-u})^{-1}[1+\gamma(u/3+8/u)$$
$$+\delta(e^u-1)^{-1}+2x^*u(e^u-1)^{-2}]. \quad (24)$$

The relations corresponding to Eqs. (20) to (23) for the transitions $j-1\rightarrow j$ are obtained by replacing j by $-j$. Because of this symmetry, the sum of terms from the transitions $j\rightarrow j-1$ and $j-1\rightarrow j$ for every j is a function of j^2 only, and the linear terms in j will cancel. Hence the following result is obtained from Eq. (19), Eqs. (2a), (20) to (23), and Eq. (2b) together with the relations corresponding to Eqs. (20) to (23) for the transitions $j-1\rightarrow j$:

$$E=30u^5\pi^{-4}e^{-u}f(\gamma,\delta,x^*,u)[(\gamma b/\omega^*)(AX/\omega^*)]^{\frac{1}{2}}. \quad (25)$$
$$\sum_j 2j^{\frac{1}{2}}e^{-\gamma u j^2/2}[1+\gamma^2 j^2 g(\gamma,\delta,u)],$$

where the temperature-independent quantity A is given by the relation

$$A=(N_T\pi\epsilon^2/3\mu c^2 u)(\omega^*/\omega_e). \quad (26)$$

The function $f(\gamma,\delta,x^*,u)$, which is obtained from Eq. (24),

$$f(\gamma,\delta,x^*,u)=1-\gamma(u/6+4/u)$$
$$-\delta[2(e^u-1)]^{-1}-x^*u(e^u-1)^{-2}, \quad (27)$$

is evidently close to unity. The function $g(\gamma,\delta,u)$ is

$$g(\gamma,\delta,u)=(3/2)(ue^{-u})^2(1-e^{-u})^{-2}$$
$$+ue^{-u}(1-e^{-u})^{-1}(5u/2+\delta/2\gamma-3)$$
$$+u(\delta/\gamma-9/2+9u/8)+27/2-7\delta/2\gamma. \quad (28)$$

The Euler-Maclaurin summation formula[22] can now be employed to evaluate the summation over j appearing in Eq. (25). Details are given in the Appendix. The result is

$$E=30u^5\pi^{-4}e^{-u}[(\gamma b/\omega^*)(AX/\omega^*)]^{\frac{1}{2}}$$
$$\times f(\gamma,\delta,x^*,u)\{1.225(2/\gamma u)^{\frac{1}{2}}$$
$$\times[1+(3\gamma/2u)g(\gamma,\delta,u)]-0.417\}. \quad (29)$$

Since $f(\gamma,\delta,x^*,u)$ is close to unity and, at room temperature,

$$(3\gamma/2u)g(\gamma,\delta,u)\ll 1, \quad 0.417\ll 1.255(2/\gamma u)^{\frac{1}{2}},$$

the following close approximation is obtained:

$$E=0.6345u^5e^{-u}(\gamma u)^{-\frac{1}{2}}[(\gamma b/\omega^*)(AX/\omega^*)]^{\frac{1}{2}}. \quad (30)$$

Utilizing the spectroscopic constants of Sponer[14] for CO and setting[12]

$$A=22.95 \text{ cm}^{-2} \text{ atmos}^{-1},$$

it is found from Eq. (29) for $b=0.08$ cm^{-1} that

$$E=1.67\times 10^{-3}(X)^{\frac{1}{2}}$$
$$(X \text{ in cm-atmos}, b=0.08 \text{ cm}^{-1}). \quad (16c)$$

The expression for E given in Eq. (16c) is seen to be in excellent agreement with the result of numerical calculations given in Eq. (16b).[23] Hence, we conclude that Eqs. (29) and (30) represent useful approximations for nonoverlapping rotational lines under appropriate conditions. It is evident from the previous discussion that Eq. (16c) cannot apply, even approximately, for values of X much smaller than 0.25 cm-atmos. This conclusion is in agreement with the well-known fact that the absorption of radiation is not proportional to the square root of the optical density for very small values of the optical density.

The expression given for E in Eq. (29) represents a significant improvement over Eq. (14) (which was derived from Elsasser's treatment for equally spaced and equally intense rotational lines). Thus the parameters appearing in Eq. (29) are, in most cases, well-known molecular constants. However, Eq. (29) must be used with discretion, particularly as regards the calculation of emissivity as a function of temperature.[24]

APPENDIX

Application of the Euler-Maclaurin summation formula[22] to the sum appearing in Eq. (25) leads to the relation

$$2\sum_{j=1}^{\infty} j^{\frac{1}{2}}e^{-\gamma u j^2/2}[1+\gamma^2 j^2 g(\gamma,\delta,u)]\simeq 2\int_1^{\infty} j^{\frac{1}{2}}e^{-\gamma u j^2/2}$$
$$\times[1+\gamma^2 j^2 g(\gamma,\delta,u)]dj+(11/12). \quad (A1)$$

The integral from 1 to ∞ can be expressed as the difference between the integral from 0 to ∞ and from 0 to 1. The integral from 0 to 1 can be evaluated approximately as 4/3. Therefore,

$$2\sum_{j=1}^{\infty} j^{\frac{1}{2}}e^{-\gamma u j^2/2}[1+\gamma^2 j^2 g(\gamma,\delta,u)]$$
$$\simeq 2\int_0^{\infty} j^{\frac{1}{2}}e^{-\gamma u j^2/2}[1+\gamma^2 j^2 g(\gamma,\delta,u)]dj-5/12$$
$$=\Gamma(3/4)(2/\gamma u)^{\frac{1}{2}}[1+(3\gamma/2u)g(\gamma,\delta,u)]-5/12. \quad (A2).$$

[22] See, for example, reference 20, p. 431.

[23] The fact that the present analysis leads to emissivities which are somewhat too large is a necessary consequence of the use of the asymptotic relation given in Eq. (18), which overestimates $f(x)$ for every finite value of x.

[24] Compare Section IV of reference 9.

Secondary Electron Emission from Metals under Positive Ion Bombardment in High Extractive Fields

E. W. Webster, R. J. Van de Graaff, and J. G. Trump
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received September 24, 1951)

In the further study of breakdown mechanisms in high vacuum, experiments show that secondary electron emission from metals under positive hydrogen ion bombardment is small and increases only slowly with the electric field strength at the bombarded surface. Although more significant emission may yet be found with the heavier ions obtained in an actual vacuum gap, the particle interchange component of the breakdown process appears to be quantitatively inadequate on the basis of present measured electron and positive ion emission coefficients. The possible importance of negative ion emission deserves consideration in the particle exchange process.

INTRODUCTION

THE work described in this report is part of a program to study the fundamental mechanisms of electrical breakdown between metal surfaces in high vacuum. The vacuum breakdown phenomenon has been reviewed by Trump and Van de Graaff[1] and the status of the electron-ion-photon exchange theory[2] has been assessed. This theory may be stated as follows: An electron in the interelectrode space is accelerated towards the anode and on impact liberates A positive ions and C photons which travel to the cathode. On arrival each ion will liberate B electrons and each photon D electrons so that the number of electrons in the interelectrode space is now $(AB+CD)$. These coefficients are dependent primarily on the total gap voltage, the electric field at the metal surfaces, and the nature of the metal and its contaminants. If $(AB+CD)>1$, the number of electrons will multiply and breakdown will ensue. This theory satisfies the observation that positive ions of the anode material take part in the reaction,[3] and can explain the dependence of breakdown upon the total voltage between the electrodes rather than on the gradient at the cathode.

The coefficient A has been measured[1] as a function of electron energy for the case of a parallel plane steel gap, the ions being subjected to an assisting extracting electric field. This coefficient increases sharply near the breakdown voltage to about 2×10^{-3}. Hence if this AB mechanism is to be important, either the value of coefficient B must be of the order of several hundreds near the breakdown voltage, or some other particle process must be present.

EARLY WORK

Experiments conducted to determine coefficient B, i.e., number of secondary electrons emitted per positive ion from a metal surface under bombardment, by Allen,[4] Linford,[5] Hill *et al.*,[6] and Trump and Van de Graaff,[1] using in all cases a low extractive cathode field of the order of tens of volts per cm, have yielded a value from 2 to 20, the number increasing with incident ion energy. The coefficient B obtained for these small collecting fields appears to reach a limiting value at relatively low voltage gradients. It had been suggested by Trump and Van de Graaff that copious secondary electron emission might be produced in the presence of high electric fields by the impact of the relatively heavy positive ions released in actual vacuum gaps from the positive electrode and its superficial layers of absorbed gas and condensed organic vapors. Such ions would distribute a portion of the energy acquired in the interelectrode gap to the superficial layers of cathode contamination and absorbed gas and to the underlying depths from which escape of secondaries is still possible. The secondary electron emission might be expected to be strongly dependent on the molecular weight and size of the impinging ion, since these would influence the area from which emission would occur as well as the energy imparted to the emission volume. It was further suggested that the low values of coefficient B previously observed might be accounted for by space charge limitation acting during the short emission time associated with the passage of the ion through the emission depth and the subsequent dissipation of its energy by the rapid recombination and high conductivity characteristic of metal surfaces. In this event a large exponential increase of secondary electrons with electric field would be expected. Although the curves of both Hill *et al.* and Linford showed a saturation in electron emission at low collecting voltages and no sign of the $\frac{3}{2}$ exponential space charge law, it appeared reasonable that higher electric fields approaching those producing field emission might produce greatly increased current flow. The availability of a source of analyzed hydrogen ions with energies up to 250 kev afforded the opportunity of measuring coefficient B under conditions of high cathode gradient.

[1] J. G. Trump and R. J. Van de Graaff, J. Appl. Phys. **18**, 327 (1947).
[2] Van Atta, Van de Graaff, and Barton, Phys. Rev. **43**, 158 (1933).
[3] Anderson, Elec. Eng. **54**, 1315 (1935).
[4] J. S. Allen, Phys. Rev. **55**, 236 (1939).
[5] L. H. Linford, Phys. Rev. **47**, 279 (1935).
[6] Hill, Buechner, Clark, and Fisk, Phys. Rev. **55**, 463 (1939).

EXPERIMENTAL ARRANGEMENT

Several methods of obtaining high cathode gradients at relatively low collecting voltages were considered in order to simplify the problem of insulating two field-producing electrodes without excessive leakage between them. A sphere-plane gap with a maximum gradient of three times that existing at a parallel plane through the sphere diameter was compared with a wire and concentric cylinder arrangement. The latter was chosen since it involved a simpler system of location and insulation of the cathode with respect to the anode, since wires of various metals are more easily procured and mounted than spheres, and since with suitable ratios of radii, higher concentrations of gradient are possible. For a given collecting voltage V across this arrangement, a gradient

$$E = V/r \log_e R/r$$

exists at the wire, where R and r are the cylinder and wire radii, respectively. For example, with $r = 0.1$ cm and $R = 1.0$ cm, $E = 4.35$V.

A section through the cylindrical bombarding chamber is shown in the drawing of the over-all experimental arrangement depicted in Fig. 1. A beam of positive ions was projected into the cylinder and impinged on the wire which was at a variable negative potential with respect to the cylinder. Those ions which did not strike the wire were collected in a Faraday cage immediately beyond the wire; the ion current to this cage was measured merely to monitor the beam and to assist its alignment with the wire. A quartz disk sealed on the bottom of the cage also facilitated positioning of the beam with respect to the wire. Both the secondary electron current from the wire to the cylinder and the ion current to the wire were measured on microammeters readable to 0.01 μA as shown in the circuit diagram of Fig. 1.

Precautions were taken by collimation to insure that none of the positive ion beam impinged on the cylinder and by the use of repressing voltage to insure that none of the secondaries from the wire escaped collection by the cylinder. Bias voltages were also applied to prevent collection of secondary electrons from the Faraday cage by the cylinder, and to avoid secondary electrons proceeded from the collimation diaphragm into the cylinder. By observing the meter currents both with and without the ion bombardment, any electrical leakage across the insulators of the wire, corona currents from connections to the cathode, back leakage through the high voltage power pack, or field emission currents from the wire were excluded in the secondary electron current e. Throughout the experiment, clean high-polished steel wires were used, out-gassing being effected both by ion bombardment and by heating electrically *in vacuo* to red heat. The vacuum was maintained at about 10^{-5} mm Hg.

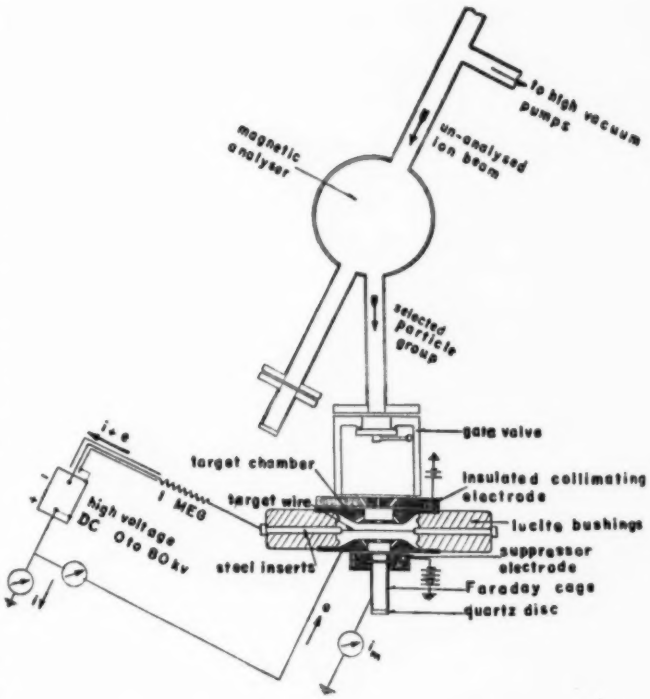


FIG. 1. Over-all experimental arrangement.

The gradient at the wire surface was varied both by change of collection voltage and by change of wire diameter. The latter was limited by the smallness of the proton current measurable and the former by the incidence of large leakage currents or even sparkover across the bushings. To facilitate changes of cathode wire without serious interference with the vacuum system a flap-type high vacuum valve was included in the pumping line immediately above the target chamber.

The positive ions were obtained from a capillary arc source using hydrogen and were accelerated to about 250 kv by means of an air-insulated electrostatic generator. Separation of protons, singly-charged hydrogen molecules, mass 3 ions, and ions of intermediate energy was accomplished by a magnetic analyzer. The analyzed ion beam currents used in this work were in the range 0.1 to 0.5 microamperes after collimation. The accelerating voltage was stabilized by a corona loading device, but this was insufficient to keep the ion current impinging on the wire cathode steady enough to permit accurate simultaneous measurements of i and e. Hence an averaging device was adopted whereby condensers were connected in parallel with the meters so as to give the same discharge time constants of several seconds for the e and i circuits. Many pairs of simultaneous current readings at practically steady state conditions were recorded for each value of extracting gradient. The coefficient B = ratio e/i was determined from each pair of readings.

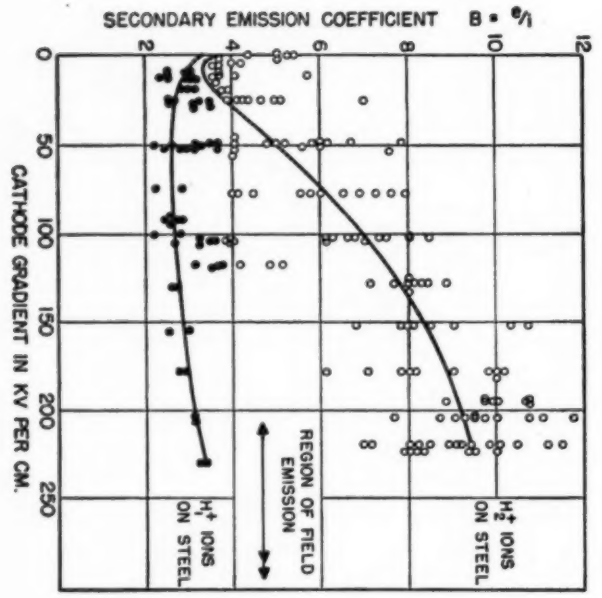


FIG. 2.

RESULTS

In Fig. 2 are plotted the series of values obtained for coefficient B for extracting gradients ranging from 0 to 200 kv/cm. Steel wires and both H_1^+ ions and H_2^+ ions were used. 200 kv/cm is the lowest gradient at which field currents were observed from the wires. The considerable spread among the individual measurements is believed to arise from the small movements of the positive ion beam and resultant variations in the angle of incidence on the target wire.

The results show a definite but slow increase in coefficient B with extracting field strength. This increase is markedly greater for H_2^+ ions than for H_1^+ ions, but never exceeds a factor of 3 and gives no indication of the large exponential dependence on field strength which had been anticipated. The coefficient B is significantly larger for the molecular hydrogen ion.

In the ion-exchange theory of electrical breakdown in high vacuum, it appears from these measurements with very light ions that field-enhanced emission of electrons by positive ion bombardment may not account in a primary way for the observed breakdown phenomena in single gaps. Further, more refined experiments are planned to extend these results. The emission properties of heavier ions such as those of nitrogen, mercury, and oil will be investigated since these are much more representative of the actual ionic particles in the vacuum gap. The possibility that accompanying negative ion emission under positive ion bombardment as observed by McKibben[7] may be the important link in this interchange will be studied.

ACKNOWLEDGMENTS

Acknowledgment is made to Charles H. Goldie, Edward Gardiner, and Robert W. Cloud for making the analyzed source of positive ions available for this experiment. This work was supported in part by the Office of Naval Research and by the grant of the Robert Blair Fellowship from the London County Council to one of the authors (EWW).

[7] McKibben and Boyer, Bull. Am. Phys. Soc. (February, 1951).

The Measurement of Ultrasonic Attenuation in Solids by the Pulse Technique and Some Results in Steel*

R. L. RODERICK† AND ROHN TRUELL
Brown University, Providence, Rhode Island
(Received October 15, 1951)

Pulse techniques for the measurement of attenuation in solids have been extended and refined sufficiently to obtain dependable measurements over a frequency range from 5 to 50 megacycles. Understanding of the relative importance of beam spreading, geometrical boundaries, and method of coupling has been improved. Coupling by means of water buffer and direct mounting is discussed in detail.

Attenuation measurements in the frequency range from 5 to 50 megacycles have been made on chrome molybdenum steel specimens, and these measurements have shown large differences in ultrasonic attenuation for samples of the same chemical composition but different heat treatment. The resulting differences are connected with anisotropy which appears in the photomicrographs. The attenuation measurements are quite sensitive to heat treatment and other factors. The application of the methods to metallurgical problems and the physics of solids is suggested.

INTRODUCTION

PULSED ultrasonic radiation for attenuation measurements in solids has been used by relatively few workers[1–5] although wide use has been made of pulsed ultrasonic energy for velocity measurements to obtain elastic moduli especially to determine the moduli of single crystals. The pulse method has been very widely used to measure both velocity and absorption in liquids and gases. The techniques used in liquids and gases are well known and have been refined to obtain sufficient accuracy for most purposes.

The measurement of attenuation of ultrasound in solids is beset with a number of difficulties both in technique and in interpretation of results, and dependable attenuation measurements in solids are very scarce. Of the two or three sets of reliable results the limitations of the methods as regards sample size, geometry, usable frequency range, and attenuation range are usually severe. The work described here has removed some of these limitations and offers methods of wider applicability and reliability than have been available previously.

It has been determined for a large class of measurements of attenuation in solids that several differences exist between true attenuation of the solid material alone and the measured attenuation of that material before allowance is made for the nature of the radiation pattern of the sound beam, the effect of the geometry of the sample, and the effect of the boundary conditions of the sample. The main source of error has turned out to be that resulting from the curvature of the wave front in the medium or what may be called beam spreading. The apparent loss because of this effect is large in certain frequency ranges, and in some cases it is the only loss measured. It is demonstrated that losses of the type mentioned can be experimentally detected and that corrections can be made for such losses when present. Certain developments in instrumentation have permitted an increase in the latitude of attenuation measurements as well as improved accuracy. A comparison has been made between the results obtained on one hand by using a water buffer between the transducer and the specimen and on the other hand by using direct coupling of transducer to specimen through a very thin cement film. It is not always easy to make such a thin film, and the water buffer is usually more reliable. If transverse waves are necessary, the water buffer obviously cannot be used and direct coupling is necessary.

The techniques developed are applied to steel specimens which are identical except for heat treatment.

Energy losses in the mechanical vibration of solids have been studied by other techniques at frequencies up to about 100 kc.[6,7]

The term "energy loss," as used here, includes energy lost from the beam by scattering and energy lost by absorption or internal dissipation in the material. Studies at the lower frequencies indicate that most solids exhibit an elastic hysteresis such that the absorption coefficient appears to be a linear function of the frequency.[8] The elastic hysteresis loss exists in many solids even at the megacycle frequencies, and it appears to be the primary loss in glasses and fused quartz.[1]

At megacycle frequencies the only method available at present for obtaining energy losses in solids is the pulse-echo technique. Rough estimates of energy loss in solids have been made by Carlin,[9] Arenberg,[10] and

* This work was supported by Watertown Arsenal under contract DA-19-020-ORD-13. The equipment used in this work was developed under this and previous contracts W-19-020-ORD-6487 and DA-19-020-ORD-11 with Watertown Arsenal.

† Now at Specialties, Inc., Syosset, Long Island, New York.

[1] W. P. Mason and H. J. McSkimin, J. Acoust. Soc. Am. **19**, 464 (1947).

[2] W. Roth, J. Appl. Phys. **19**, 901 (1948).

[3] W. P. Mason and H. J. McSkimin, J. Appl. Phys. **19**, 940 (1948).

[4] W. J. Price, Phys. Rev. **75**, 946 (1949).

[5] R. K. Roney, Ph.D. thesis, California Institute of Technology (1950).

[6] S. Siegel, J. Acoust. Soc. Am. **16**, 26 (1944).

[7] C. Kittel, Reports on Progress in Physics (The Physical Society) **11**, 205 (1947).

[8] R. L. Wegel and H. Walther, Physics **6**, 141 (1935).

[9] B. Carlin, Product Engineering, October, 1947.

[10] D. L. Arenberg, J. Acoust. Soc. Am. **20**, 1 (1948).

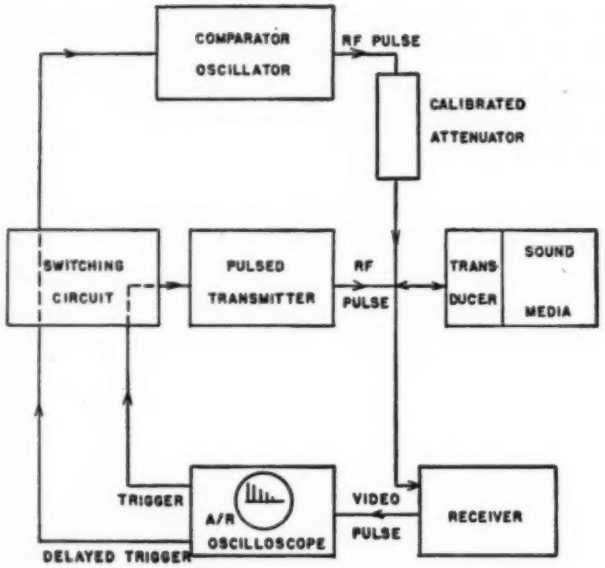


FIG. 1. Block diagram of ultrasonic pulse system.

Huntington[11] using the pulse technique. The first reliable measurements were made by Mason and McSkimin on polycrystalline aluminum rods, glasses, and fused quartz. Additional measurements on polycrystalline magnesium and aluminum have been made by Mason and McSkimin, Roth, and Roney. The explanations offered by Roney show that all of the measurements can be correlated fairly well by the attenuation expression

$$\alpha D = B_1\mu + \frac{2}{\mu^2}\left(\frac{\Delta K}{K}\right)^2 \sum_{m=0}^{\infty} (2m+1)\sin^2\delta_m,$$

where $\mu = \pi D/\lambda$, λ is the wavelength in the material, α is the attenuation coefficient, D is the average grain size, $B_1\mu$ is the linear elastic hysteresis absorption term obtained from low frequency data, $(\Delta K/K)^2$ is the square of the mean value of the fractional variation of the elastic constants of a single crystal (the elastic anisotropy factor), δ_m is defined by

$$\tan\delta_m = \frac{mj_{m-1}(\mu) - (m+1)j_{m+1}(\mu)}{mn_{m-1}(\mu) - (m+1)n_{m+1}(\mu)},$$

and j_m and n_m are the Bessel and Neumann functions. The second term is due to the "scattering" of the sound by the grains of the polycrystalline metals. When $\mu \to 0$, the scattering term is proportional to μ^4 and was given by Mason and McSkimin. When $\mu \to \infty$, the scattering term is independent of μ and was empirically obtained by Roth.

Two methods have been used here to measure the attenuation. The first method is to directly mount the piezoelectric quartz crystal on the face of the specimen.[4,10,11] The opposite face is made accurately parallel to the first. The dimensions of the specimen are made so that the reflections from the sides do not interfere with the main echoes. The second method involves the use of water as a buffer between the specimen and the transducer.[2,3,5] The sound pulse is radiated from the transducer and traverses the water path until it encounters one of the parallel faces of the specimen. Part of the energy is then reflected back toward the transducer and part of it enters the specimen. The water path is made long enough so that the sound can travel back and forth in the specimen many times before the sound makes one round trip in the water buffer; a number of specimen echoes can then be observed in the same space of time as that required for one trip from the transducer to the specimen and back. One advantage of the water buffer method is that the reflection coefficients at the specimen faces can be determined experimentally. The transmission loss is found by taking attenuation measurements first with one face submerged in the buffer and then with both faces submerged.

In comparison with attenuation measurements in solids by the pulse technique those in liquids can be obtained rather easily and to a high degree of accuracy. Pinkerton[12] gives an excellent analysis of the technique for liquids. The essential difference between liquid and solid measurements is that in liquids the path length may easily be changed, making it possible to eliminate the evaluation of reflection losses of any kind and at the same time making it possible to use only one reflected pulse. In solids it is difficult to change the path length and this change cannot be made continuously. In some cases such as with single crystals the sample must not be destroyed. Also it would be difficult to repeat measurements on the same sample at some later date. Measurement of attenuation as a function of temperature, magnetic saturation and frequency could not be done directly and in any event would be troublesome if the sample length were changed to determine the attenuation.

The availability of specimens makes it desirable that it be possible to make measurements on reasonably small specimens. The general requirement on the method of determining the attenuation is that it must work with small specimens which may not be destroyed. Since the sample size is usually fixed, it is desirable for purposes of interpretation of the data that the design of the equipment be directed toward making measurements over as large an attenuation and frequency range as possible.

EXPERIMENTAL PROCEDURE AND TECHNIQUE (DIRECT MOUNT)

The block diagram of Fig. 1 outlines the system used for the measurements; when a water buffer is used, the water buffer and the sample are the sound media. If at

[11] H. B. Huntington, Phys. Rev. **72**, 321 (1947).

[12] J. M. M. Pinkerton, Proc. Phys. Soc. (London) **B62**, 286 (1949).

first the switching circuit and the comparator oscillator are ignored, the sequence of operations is as follows. The trigger pulse from the A/R scope (about 300 pulses per second) triggers the transmitter. The transmitter produces a one-to-four microsecond pulse of radio frequency energy between 5 and 50 megacycles. This high frequency electrical energy is applied to a piezoelectric quartz crystal which converts the electrical oscillations into mechanical vibrations. These ultrasonic waves pass into the media to be tested and are returned to the transducer by reflection from the far end of the specimen and reconverted to electrical signals. By the time the returned signal has arrived the transmitter is off; the received signals enter the input of a sensitive receiver which presents an amplified video pulse to the vertical plates of the A/R scope, the horizontal sweep of which was initiated with the trigger pulse. There will, in general, be many reflected echoes arising from the first pulse which is reflected back and forth within the sample media, and many pulses will appear on the A/R scope after the initial pulse from the transmitter. The next trigger pulse from the A/R scope causes the process to be repeated.

The time between successive transmitter pulses is about 3000 microseconds, and the transit time for one round trip in the sound medium is approximately 5 to 50 microseconds. There are no standing waves in the sound medium under proper operating conditions.

A crystal calibrated time delay is available on the A/R scope which makes it possible to pick out 4, 10, and 25 microsecond sections of the full sweep at any time delay up to 1000 microseconds. The delayed trigger can be initiated at any time up to 1000 microseconds after the main trigger going to the transmitter.

The time between successive echoes gives the velocity if the length of the sound media is known. The attenuation is determined by introducing an auxiliary radio frequency pulse from another pulsed transmitter called the comparator oscillator, operated at the same frequency as the transmitter. The variable delayed trigger which operates the comparator oscillator makes it possible to delay this pulse in relation to the radio frequency pulse from the transmitter and so place the comparator pulse near the echo to be measured. The relative amplitude of each echo is determined by matching the pulse from the pulse comparator to it by inserting attenuation by means of a calibrated attenuator box and recording the attenuator readings. These readings give the relative amplitudes of the reflected pulses in decibels.

Low level anomalous reflections frequently appear between the main echoes. In addition one often uses small specimens which crowd the main echoes together. In both cases the amplitude of the comparator pulse becomes indeterminate when placed between two main echoes for comparison. In order to avoid this a switching circuit was designed which alternately pulses the transmitter and the pulse comparator so that on the A/R scope two superimposed sweeps appear, one containing the echo pattern and the other the pulse from the pulse comparator.

In the direct mounting or coupling method one cements, with a very thin film, a piezoelectric crystal (e.g., a circular 10-mc X-cut quartz crystal having a diameter of $\frac{3}{4}$ inch and a thickness of 0.0114 inch) onto a smooth flat surface of the solid specimen as in Fig. 2.

The actual measurement technique is as follows. The pulse transmitter is tuned to the resonant frequency of the quartz crystal or its harmonic. Three $\frac{3}{4}$-inch X-cut quartz crystals are used; a 5-mc crystal operated at 5, 15, 25, 35, and 45 mc, a 10-mc crystal operated at 10, 30, and 50 mc, and a 20-mc crystal operated on its fundamental. The time difference between the first and last echo within the range of the 0 to 100 microsecond time delay is recorded from which the dilatational velocity is computed. The amplitude of each echo is then compared with the pulse from the pulse comparator and attenuator readings are taken. A plot is made of attenuation as a function of the echo number and the slope is determined graphically. This, after taking into account the length of the sample, gives the apparent attenuation in decibels/unit length. The relation between these measurements and the attenuation coefficient can be found where the attenuation coefficient, α, is defined as

$$\alpha=\lim_{\Delta z\to 0}\frac{\sigma(z)-\sigma(z+\Delta z)}{\sigma(z)}\Big/\Delta z=-\frac{\sigma'(z)}{\sigma(z)},$$

where $\sigma(z)$ is the maximum stress amplitude of a dilatational or a rotational plane sinusoidal wave traveling in the z-direction. $\sigma(z)$ can be either a normal or a shear stress. Then

$$\sigma(z)=(\text{const.})e^{-\alpha z},$$

where α is assumed to be the same in all parts of the solid through which the wave travels and independent of the stress amplitude. The actual measurements, taken in decibels/unit length, are related to α in nepers per unit length by

$$\alpha\left(\frac{\text{decibels}}{\text{unit length}}\right)=8.686\alpha\left(\frac{\text{nepers}}{\text{unit length}}\right).$$

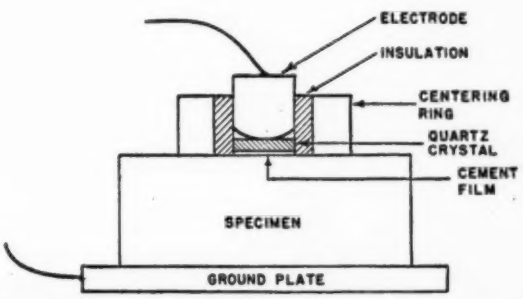


FIG. 2. Electrode arrangement with crystal mounted directly on specimen.

It has been assumed that the electrical output of the quartz transducer is directly proportional to the amplitude of stress applied to it from the solid.

Since the measurements were made on conducting specimens, the specimen was taken as one of the electrodes. The other electrode is a piece of brass pressed against the crystal by its own weight. No great care had to be taken with the contact between this electrode and the crystal since the dielectric constant of quartz is reasonably small. Figure 2 shows the electrode and the centering ring in position. Rounding the electrode at the edges tends to reduce the side lobes in the radiation pattern. Oils, Vistac, phenyl salicylate, and phenyl benzoate have been tried as cement films between the quartz crystal and the specimen. In every case repeatable results were secured by "wringing" the crystal onto the specimen much as one does with gauge blocks. Reasonably good transmission was obtained with all. The oil films changed very rapidly with time. The Vistac gave higher losses at the higher frequencies than the phenyls. The phenyl benzoate has a higher melting point (70°C)

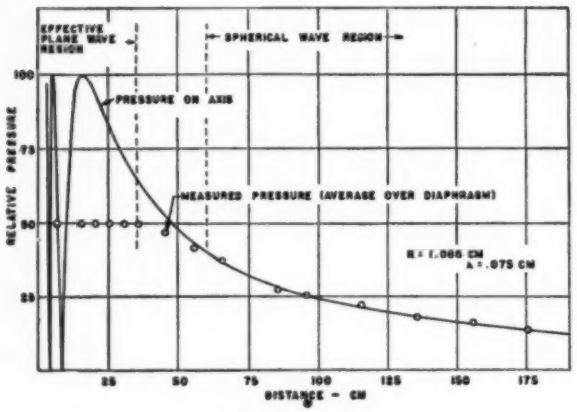


FIG. 3. Sound field of circular piston radiator.

than salol (43°C) requiring the sample to be raised to a higher temperature, but phenyl benzoate was easier to use because it solidified more rapidly. Generally, however, salol has been used here. The Vistac and oils caused poorer pulse shapes. The use of the "wringing in" technique with salol gave results which were repeatable to within a few percent on remounting the crystal.

The crystal when excited electrically produces a radiation pattern of sound waves traveling into the specimen. Some of the energy is lost through beam spreading, and it is necessary to determine either experimentally or analytically the nature of this pattern in order to determine the corrections to be applied to the attenuation data. In some cases errors of 1000 percent or worse can be traced to this source.

The first approach to the analysis of this problem is given by the extensive work that has been done both experimentally and analytically on the radiation pattern of a circular piston source in an infinite rigid baffle radiating into a half-space filled with fluid. The excited piezoelectric crystal is made equivalent to an oscillating piston-like displacement. The amplitude of the pressure, $p(z)$, on the axis of symmetry can be found explicitly as

$$p(z)=p(0)\sin\{\tfrac{1}{2}k[(z^2+a^2)^{\frac{1}{2}}-z]\},$$

where z is the distance along the axis of symmetry, $k=2\pi/\lambda$, λ is the wavelength in the medium, and a is the radius of the piston source. Near the sound source the waves are nearly plane and the on-axis pressure, given by the foregoing formula has a series of maxima and minima because of the interference of the sound waves. Figure 3 taken from Simmons and Urick[13] shows a plot of this function for $a=1.086$ cm and $\lambda=0.075$ cm (in water at 2 mc). Simmons and Urick measured the amplitude of a sound pulse reflected back to the transmitting crystal by a small reflector. The position of the last maximum is at $z_0/2=a^2/\lambda$. The experimental curve in Fig. 3 shows that the effective plane wave region extends to z_0 (twice the distance from source to last maximum of the on-axis pressure amplitude) because of the off-axis character of the radiation pattern, i.e., the pressure remained constant up to twice the distance from the source to the last maximum of the on-axis pressure amplitude. The receiver used here is the radiating sound source itself which averages the received sound over its surface, and since a pulse of finite duration is used a time average occurs also. The net result is that the averaged pressure amplitude is constant in the near or plane wave region.

Far from the circular piston source the pressure amplitude, $p(r,\theta)$, is given by[14]

$$p(r,\theta)\sim(\text{const})\frac{e^{-irk}}{r}\frac{J_1(ka\sin\theta)}{ka\sin\theta},\quad r\gg a,$$

where r is the radial distance from the center of the piston source and θ is the angle from the axis of symmetry to r. Hence, the sound waves are spherical and decay like $1/r$ far from the source.

In measuring the attenuation in liquids by the pulse method Pinkerton[12] found it necessary to use a large reflector for averaging in the plane wave region and to correct his measurements for beam spreading in the spherical wave region.

If a circular crystal is cemented on the surface of a solid as is done here in the direct mounting technique, the radiation pattern differs in several ways from the radiation pattern in a fluid as discussed previously. An analysis of the sound radiation pattern obtained for a piston-like sound source on a semi-infinite solid has been made by one of the authors.[15] One finds that even though the source is vibrated longitudinally the restraining influence of the stress free surface causes shear waves to

[13] B. D. Simmons and R. L. Urick, J. Acoust. Soc. Am. **21**, 633 (1948).

[14] P. M. Morse, *Vibration and Sound* (McGraw-Hill Book Company, Inc., New York, 1948).

[15] R. L. Roderick, Metals Research Laboratory Report, Brown University (1950).

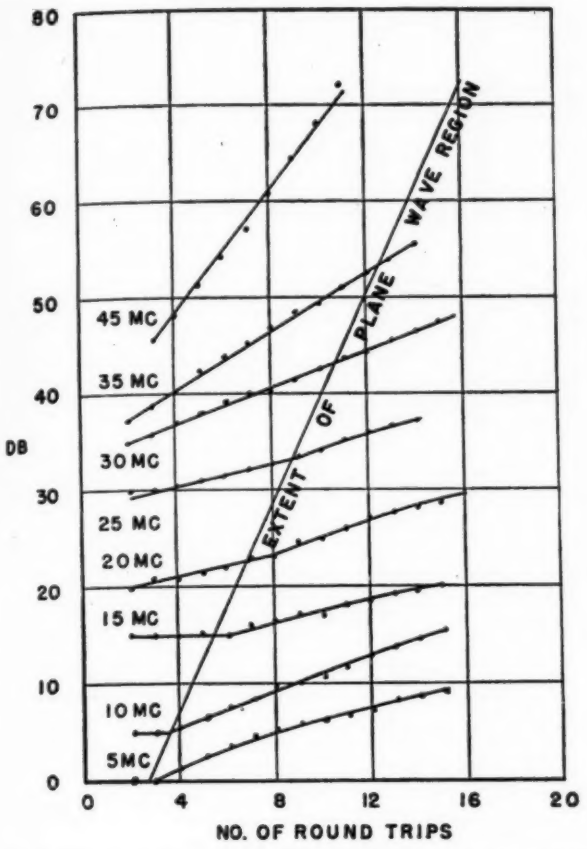


FIG. 4. Attenuation results in steel (Rockwell "C" 61) with crystal mounted directly on sample.

be produced. Far from the source the radiation pattern of the longitudinal waves is of the same general nature as the radiation pattern for the pressure waves in fluids under similar circumstances. The determination of the radiation pattern near the source would require rather lengthy numerical work.

The existence of a plane wave region and a spherical wave region for a piston source on a solid is demonstrated in Fig. 4. The relative amplitude in decibels of the series of pulses reflected back and forth in a hard chrome-molybdenum steel (Rockwell "C" 61) is plotted for frequencies from 5 to 45 mc. The loss in amplitude is plotted positively. If we assume the radiation pattern in the solid is like the fluid radiation pattern experimentally checked by Simmons and Urick,[13] then the plane wave region should extend to $2a^2/\lambda$. This point was calculated for each of the frequencies used and plotted on the attenuation curve, Fig. 4, as the extent of the plane wave region. For this calculation $a=\frac{1}{4}$ inch and $c_D=2.13\times10^5$ inch/sec, the dilatational velocity of sound in the steel. The general agreement of the theoretical points with the break in the actual attenuation curves is better than one should expect at the lower frequencies in view of the underlying assumptions that a longitudinal wave in a solid behaves like a pressure wave in a fluid, that the crystal is a piston source, and that the many reflections from the plane surfaces of the solid do not greatly alter the radiation pattern from the pattern in a semi-infinite solid. Since the 1000-volt pulse from the transmitter goes directly into the receiver and saturates it for about 10 microseconds, the first echo is unusable. In highly attenuating solids where only a few echoes can be detected the time dependence of gain resulting from saturation can be determined, by use of the comparator oscillator, with reasonable accuracy and with valid measurements made on all echoes. The time dependence of gain is determined by finding the relative amplitude of the pulse from the comparator oscillator at a series of points along the sweep.

It is possible to obtain fair response when driving a crystal at nearly any frequency because of the loading of the crystal by the solid. The operating frequency must, however, be at the natural resonant frequency of the crystal or one of its harmonics for attenuation measurement. This fact can be seen from the curves in Fig. 5. The attenuation and velocity were measured in the neighborhood of the fifth harmonic of a 5-mc X-cut quartz crystal. The minimum in attenuation and the maximum in velocity show clearly the position of the proper operating point. The explanation of the observed attenuation effect is that when the sound pulse is reflected from the solid-crystal interface, some of the energy passes into the crystal and is reflected from the crystal-air interface back into the solid. At the resonant frequency this energy comes back in phase with the

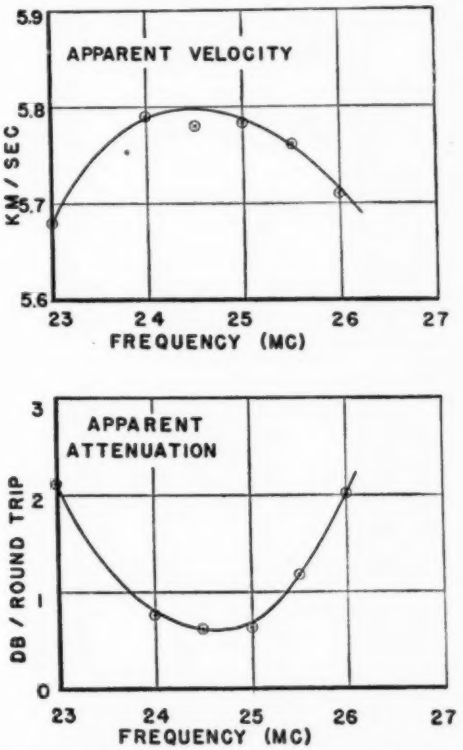


FIG. 5. Effect of operating around the resonant frequency with directly mounted crystal with steel sample (Rockwell "C" 61) and 5-megacycle X-cut crystal and salol cement.

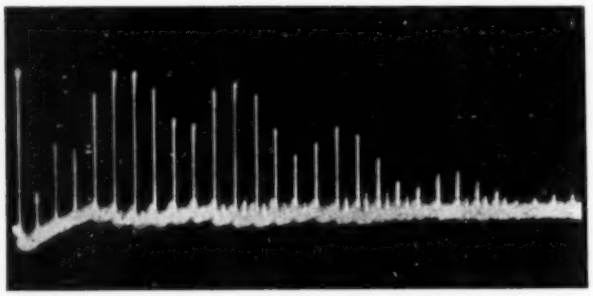


FIG. 6. Echo pattern in specimen with large ratio of sound path length to diameter for the frequency used. Small "in between" pulses arise from mode conversion at sidewalls. (Photograph retouched.)

reflected sound pulse. Off the resonant frequency the energy comes back out of phase thereby decreasing the amplitude of the echo. Substantially the same result is obtained at the other harmonics. With a thinner 10-mc X-cut crystal the effect is less pronounced. The best velocity measurements are made with a light coupling such as oil or Vistac and a thin crystal. The attenuation limits the maximum usable frequency and hence the thinness of the crystal used.

Figure 6 shows the type of echo pattern obtained when the diameter of the sample is too small compared to the path length for the sound. If the diameter of the sample is large enough, then energy reflected back from the side walls returns to the crystal at a much later time and does not stay with the proper earlier pulses. Interference effects account for the pattern observed. From Fig. 7 it is seen that the energy radiated from the crystal at an angle will return to the crystal on the nth echo where $n=R/L$. The first lobe of the radiation pattern of a piston source is at an angle $\theta=0.85\lambda/a$. From this an estimate can be made on the size of a sample for a particular measurement in order that the energy being returned by the sides of the sample does not interfere with the pulses on which the measurements are taken. In general the lowest frequency gives the most trouble since θ is inversely proportional to frequency. The interference effects just mentioned can, of course, only arise when there is sufficient beam spreading and when the path to the side walls and back is sufficiently short. If beam spreading could be eliminated completely, the effect mentioned would obviously vanish even for samples small in diameter. If the sample is sufficiently large in diameter, energy loss through beam spreading alone will occur and this will be reduced as the frequency is increased.

The effect of nonparallelism on the reflected echoes when one uses a common sending and receiving crystal can be estimated as follows. The crystal of radius a is assumed to act as a piston source on a semi-infinite stress free solid. The wave front is spherical with center at the crystal center far from the source.

Assume the sample is slightly out of parallel by an amount α and that the crystal acts as a spherical point

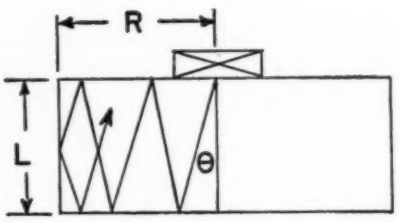


FIG. 7.

source. In Fig. 8, θ' is the angle between the line from the center of the crystal to the normal to the opposite face and the emerging ray. The distance between faces is d. It turns out that the distance, X_n, from the origin of radiation (center of crystal) to the point where the nth echo returns to the plane of the crystal can be expressed as

$$X_n=\sum_1^n \chi_i,$$

where

$$\chi_n=\frac{\sin 2[\theta'-2(n-1)\alpha]}{\cos[\theta'-(2n-1)\alpha]}\,\frac{d-\sin\alpha\sum_1^{n-1}\chi_i}{\cos[\theta'-2(n-1)\alpha]}.$$

The angles θ for which the rays return to the crystal on the nth echo can now be found from $-a\leq X_n(\theta')\leq a$. Since $n\alpha\ll 1$, the foregoing expression can be summed giving $X_n/2d=n\tan\theta'-n(n-1)(1-\tan^2\theta')\alpha$, from which θ' in the inequality could be computed numerically in every case. If also $\theta'\ll 1$ and letting $\theta'=\theta-\alpha$ then

$$-\frac{a}{2nd}+n\alpha\leq\theta\leq\frac{a}{2nd}+n\alpha.$$

This is a beam centered at $\theta_c=n\alpha$, of width a/nd, and forms the nth echo.

Actually the beam forming the nth echo is an elliptical cone. The pressure at the center of this cone is a measure of the response of the crystal to the nth echo.

A combination of these results shows that minima should occur in the echo patterns approximately when $\theta_c(n)=\theta_i$ where n=echo number, θ_i are the zeros of $J_1(ha\sin\theta)$ and $\theta_1=0.6\lambda/a$, $\theta_2=1.1\lambda/a$, and $\theta_3=1.71\lambda/a$.

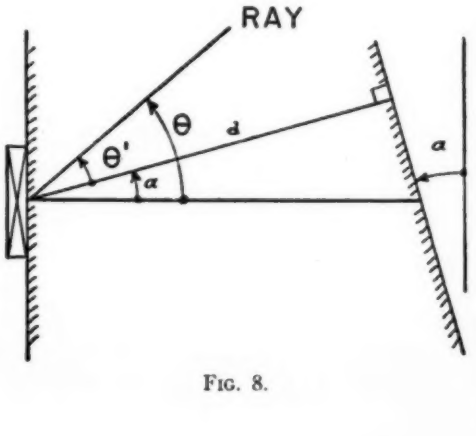


FIG. 8.

In order to be sure nonparallelism has no effect on the reflected echoes, it is necessary that $\alpha \ll 0.6\lambda/an$ where n is the number of echoes detected. Using steel and a crystal with $a=\frac{1}{4}$ inch and taking 10 as the number of usable echoes one obtains the condition $\alpha \ll 0.055/f$ where f is in megacycles. This requirement is easily met by using a surface grinder to grind the flats on the specimens.

EXPERIMENTAL PROCEDURE AND TECHNIQUE (WATER BUFFER)

The difference between the direct mounting method and the water buffer method lies in the provision of an isolation medium between the sound source (transducer) and the specimen, Fig. 9. The isolation medium (in this case, water) makes it possible to measure directly the reflection loss. Whereas in the direct mounting technique, the reflection loss cannot be measured directly and is dependent on the technique of crystal mounting and the thickness of the resulting cement film, as well as on pulse length and other factors. There is, moreover, no pulse distortion at the reflecting surfaces as is the case with the direct mounting technique. The maximum measurable attenuations are as large as those that can be made by directly mounting the transducer on the specimen. The band pass of the water loaded transducer does limit the minimum pulse length, whereas a solid loading makes possible transmission of very short pulses although these pulses will be quite distorted.

Figure 9 shows the sound path. The paths followed by the sound are drawn at an angle only in order to distinguish them. The relative amplitude of the transmitted and reflected wave at the solid-liquid interface is determined by the ratio of the acoustical impedances,

$$\zeta = \frac{(\rho c) \text{ for liquid}}{(\rho c_D) \text{ for solid}}.$$

The transmission and reflection losses are useful here when expressed in decibels (db). A loss is taken as positive. The loss in amplitude of a wave going from the liquid into the solid is

$$T_1 = 20 \log\left[\frac{(1+\zeta)}{2\zeta}\right].$$

TABLE I.

Rockwell "C"	Fused quartz	S.A.E. 4150 steel 61	43	18	14	Units
$c_D \times 10^{-5}$	5.87	5.83	5.87	5.88	5.86	cm/sec
$c_R \times 10^{-5}$	3.75	2.77	3.20	3.18	2.79	cm/sec
ρ	2.206	7.800	7.808	7.816	7.845	g/cm³
$\mu \times 10^{-11}$	3.10	5.98	7.99	7.90	6.10	dynes/cm²
$\lambda \times 10^{-11}$	1.40	14.53	10.87	11.22	14.70	dynes/cm²
$E \times 10^{-11}$	7.16	16.21	20.61	20.46	16.53	dynes/cm²
σ	0.155	0.354	0.288	0.298	0.354	
$\rho c_D \times 10^{-5}$	12.95	45.47	45.81	45.96	50.99	g/cm²/sec
T_1	13.71	23.94	24.02	24.04	24.92	db
$-T_2$	5.08	5.74	5.74	5.74	5.77	db
$R_1 = R_2$	2.00	0.57	0.56	0.56	0.51	db
R_2 meas ave	2.14	Average for all specimens			0.66	db

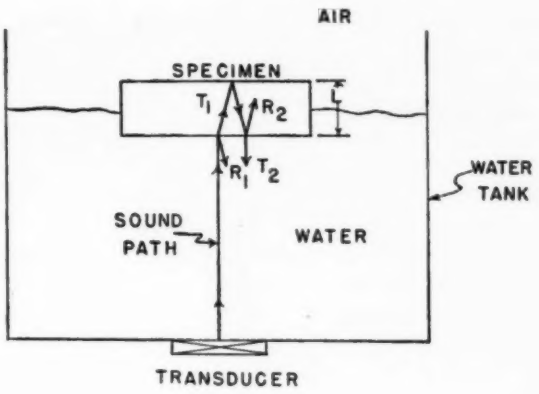


FIG. 9.

The other coefficients are

$$R_1 = 20 \log\frac{1+\zeta}{1-\zeta} = R_2 \quad \text{and} \quad T_2 = -20 \log\frac{2}{1+\zeta}.$$

These coefficients are tabulated in Table I for the steel specimens used in this investigation and for fused quartz using water as the liquid buffer (ρc for water was taken as 1.49×10^5 g/cm² sec).

The relative amplitudes of the successive reflections in the solid are measured in the same manner previously outlined. The logarithm of the ratio of the amplitude of the nth echo to the amplitude of the energy incident on the solid from the liquid is $T_1 + 2nL\alpha + (n-1)R_2 + T_2$. The slope of this expression as a function of n is taken from the measurements, i.e., $2L\alpha + R_2$. If the other face of the sample is also submerged in the water, then the expression for the slope will be $2L\alpha + 2R_2$. The difference in slopes is, then, the reflection coefficient. The average of the reflection coefficients measured for the steel samples was 0.66 decibel over a frequency range of from 5 to 50 mc with a scatter of from 0.5 to 0.8 decibel. Considering this it was felt that the value obtained from the velocity data was satisfactory, and this value was used in computing the attenuation coefficient, α, from the slope of the attenuation curve obtained with one surface of the specimen submerged. The large value of the reflection coefficient, R_2, for fused quartz (2.00 decibel) makes attenuation measurements in quartz more difficult than equally accurate measurements in steel where R_2 is smaller.

Perhaps the most important single factor in the measurement of attenuation in solids is that of curvature of wave front or beam spreading. The effect of beam spreading on the measurements is especially pronounced in the specimens of low attenuation, see Figs. 14, 15, and 16. The extent of the plane wave region and the exact nature of the beam spreading is complicated slightly by the addition of the water buffer although an estimate of the extent of the plane wave region can be made by taking into account the slower velocity of sound in water. The average slope of the measured attenuation

curves corrected for reflection losses but including the beam spreading was taken from Figs. 14 and 15 for fused quartz and steel 61 and plotted in Fig. 10. These curves of apparent attenuation *versus* frequency show the approximate magnitude of the beam spreading correction at 5 and 10 mc. The attenuation in fused quartz extrapolated from data of Rapuano[16] is shown for comparison.

Errors which can be large will arise when the sample is not precisely aligned with the quartz crystal. The procedure for alignment was to maximize the echoes from the solid water interface; this seemed to be the most sensitive and satisfactory procedure.

Figure 11 shows the water tank with a specimen in the specimen holder. The three crystal holders contain 5-, 10-, and 20-mc crystals used to cover the 5- to 50-mc frequency range. The construction details of the crystal holders were taken from Roth.[17] The specimen holder was constructed so the specimen could be fully submerged to measure the reflection coefficient without interference from reflections from the holder. The aligning head provides rotation about two perpendicular axes. The use of a milling table and a standard drill press stand makes the equipment versatile and easy to construct as well as rigid.

The maximum measurable attenuation was limited, primarily, by two factors. Cavitation in the water, especially at the lower frequencies where the transmitter was most efficient, limited the maximum amplitude of the sound pulse. The minimum detectable echo was determined by the signal-to-noise ratio of the first radio frequency amplifier.

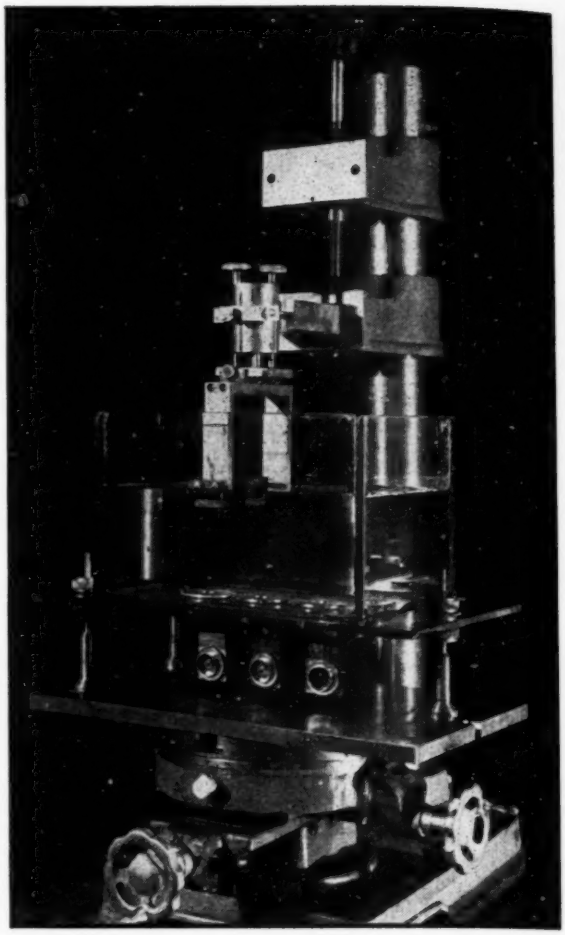


FIG. 11. Water buffer tank and mounting.

ELECTRICAL EQUIPMENT

The design considerations in constructing the pulsed transmitter and receiver for optimum performance were many and often conflicting. Those of most importance were best signal-to-noise ratio in the receiver, maximum excitation of the transducer, best receiver recovery time, and easily variable frequency. A method of measuring the attenuation was used which avoided the many difficulties arising from the high frequency used and from the nonlinearity of the receiver both with regard to time and amplitude. This equipment was designed primarily by Bronzo.[18] The block diagram, Fig. 12, shows the system in detail. Figure 13 is a photograph of the equipment set up to take measurements by the water buffer technique.

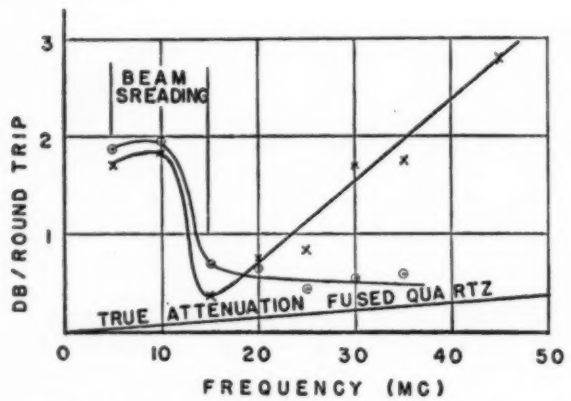


FIG. 10. Effect of beam spreading in water buffer using fused quartz (points ⊙) and steel 61 (points ×).

EXPERIMENTAL RESULTS AND DISCUSSION

VELOCITY MEASUREMENTS

The direct mounting technique was used to measure the velocity. The time difference between the first echo and the last echo within the range of the 100-microsecond time delay was measured and divided by the number of echoes; this and the length of the sample gave the velocity. A 10-mc X-cut crystal was used for measuring the dilatational velocity and 5- and 10-mc AC-cut crystals were used for measuring the rotational

[16] R. A. Rapuano, Tech. Rep. 107, Research Laboratory of Electronics, M.I.T. (1949).
[17] W. Roth, Ph.D. thesis, M.I.T. (1947).
[18] J. A. Bronzo, Metals Research Laboratory Report, Brown University, (1950).

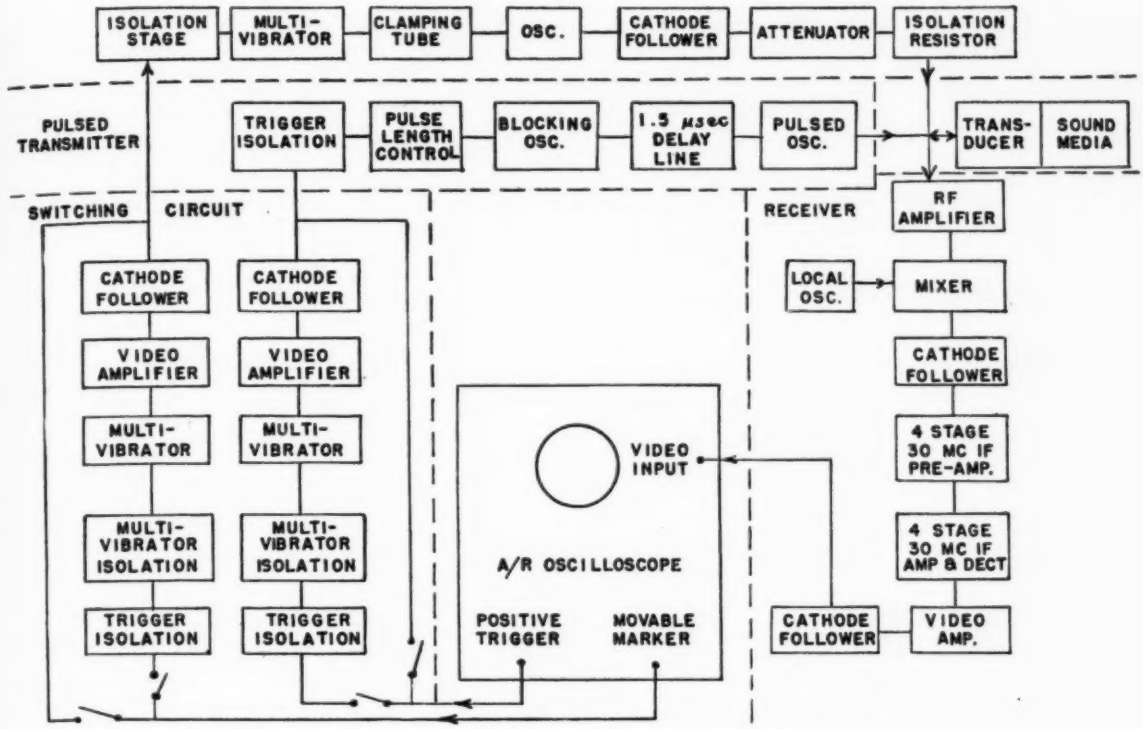


FIG. 12. Complete block diagram of ultrasonic pulse equipment.

velocity. Other measurements were made at the higher frequencies and to within the accuracy of the measurements the velocity was constant. The effect of the crystal on the effective sound path was not taken into account. The density was measured by submersion in water.

The elastic constants can be computed from the velocities and density as

$$\mu=\rho c_R^2, \qquad \lambda=\rho c_R^2(R^2-2),$$

$$E=\rho c_R^2\left(\frac{3R^2-4}{R^2-1}\right), \qquad \sigma=\frac{R^2-2}{2(R^2-1)},$$

where c_D is the dilatational velocity, c_R is the rotational velocity, $R=c_D/c_R$, λ and μ are the Lamé constants, E is Young's modulus, and σ is Poisson's ratio. The velocities, densities, and elastic constants for the steel specimens and fused quartz are tabulated in Table I. In addition the reflection and transmission coefficients are tabulated along with the acoustical impedence. The average measured value of R_2 is included. The definitions of T_1, T_2, R_1, and R_2 have been given.

ATTENUATION MEASUREMENTS

The most reliable attenuation measurements were obtained by the use of a water buffer between the specimen and the transducer. The technique for making the measurements has been described. Figures 14, 15, 16, and 17 present the relative amplitude of the successive echoes in the respective specimens for each of the frequencies used. The relative decibel loss is plotted as a function of the number of round trips and a loss in amplitude is plotted positively. The curves are arbi-

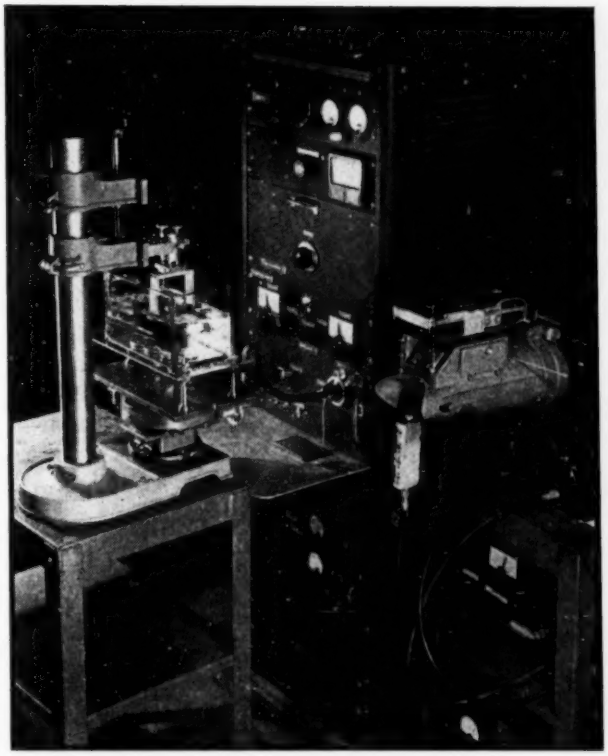


FIG. 13. View of equipment as used with water buffer.

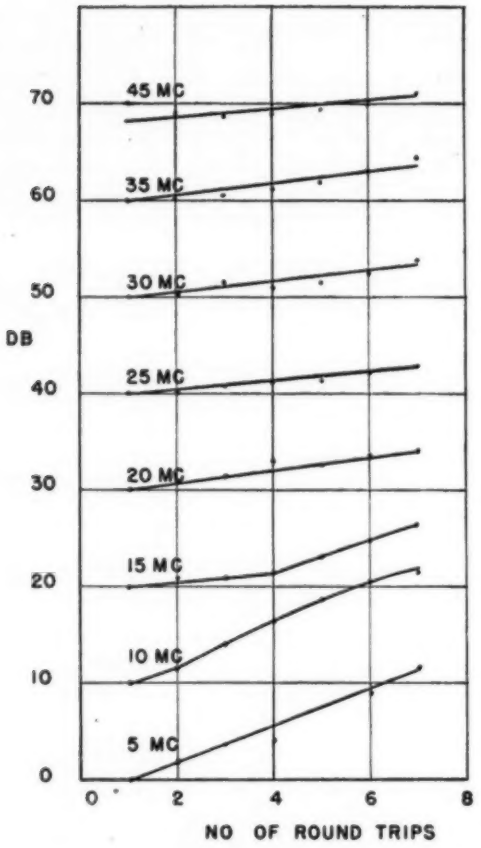


FIG. 14. Attenuation values taken on fused quartz using water buffer. Data corrected for reflection loss. Reflection loss = 2.00 db/reflection. Specimen thickness 0.7500 in.

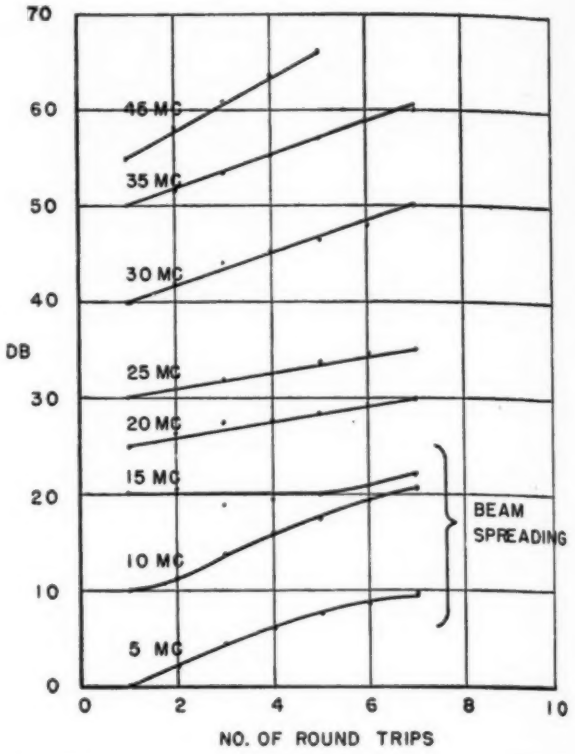


FIG. 15. Attenuation values taken on hard steel (Rockwell "C" 61). Data corrected for reflection loss. Reflection loss = 0.57 db/reflection. Specimen thickness 0.7500 in. Water buffer.

trarily separated. Only the slope of the curves has significance. In order to make all the curves comparable the reflection loss, computed from the velocity data, was subtracted from the data. The actual measurements (data) were taken with only one face of the sample submerged in the water buffer.

Since fused quartz has a very small attenuation, especially in the frequency range used here, it has been used to ascertain the sources of error, in particular to indicate the magnitude of the beam spreading and the extent of the plane wave region. The velocity in fused quartz is very nearly the same as the velocity in steel for longitudinal waves so that the quartz sample could be made the same thickness as the steel samples for direct comparison to determine the beam spreading. All samples were 0.7500 inch in thickness so that a round trip represents 1.5 inches of travel.

Figures 14, 15, and 16 for the fused quartz, steel 43, and steel 61 samples bring out clearly the extent of the beam spreading region and where valid attenuation measurements can be made on low attenuating samples. The apparent attenuation in fused quartz at higher frequencies is caused by error introduced by the relatively high percentage of the measured attenuation represented by the reflection losses and to the difficulty

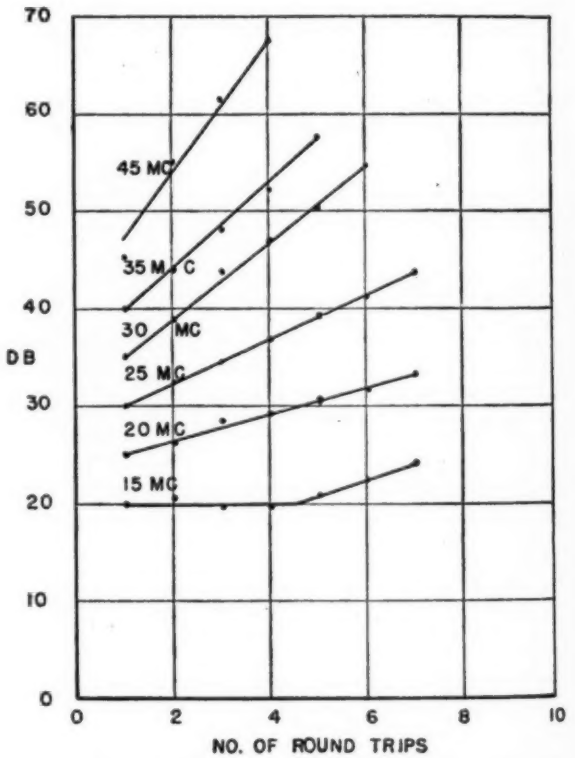


FIG. 16. Attenuation values on steel (Rockwell "C" 43). Data corrected for reflection loss. Reflection loss = 0.56 db/reflection. Specimen thickness 0.7500 in. Water buffer.

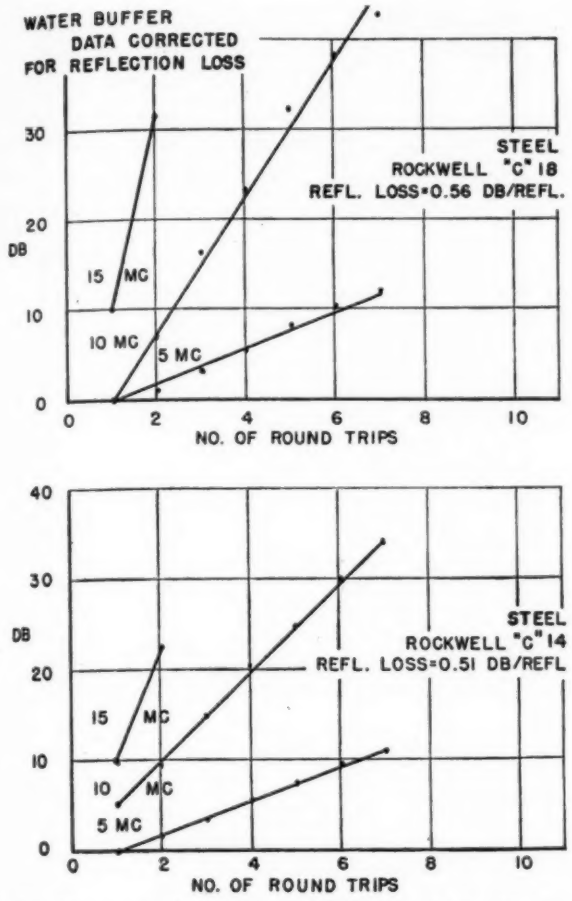


FIG. 17. Attenuation values on steels 14 and 18 using water buffer. Data corrected for reflection loss as indicated. Specimen thickness 0.7500 inch.

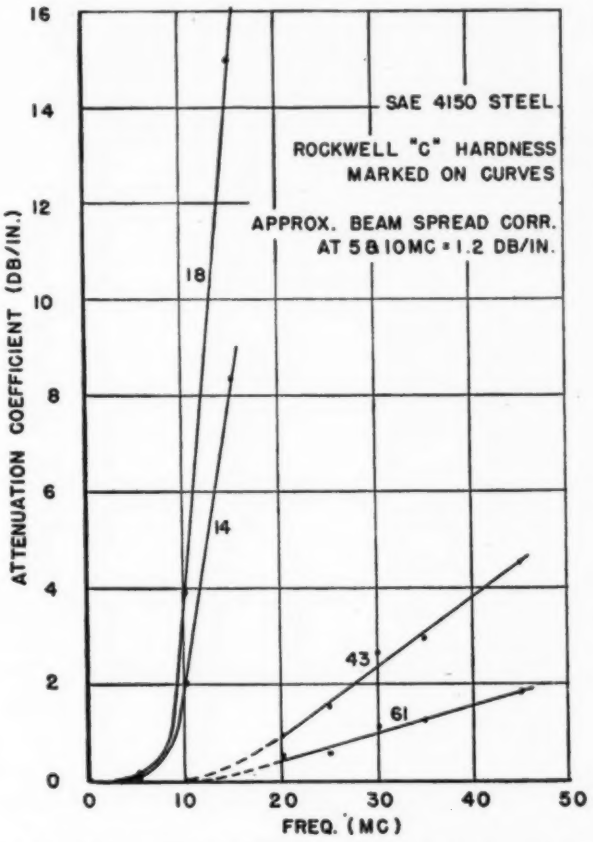


FIG. 18. Attenuation-frequency plot for four steel samples. Water buffer.

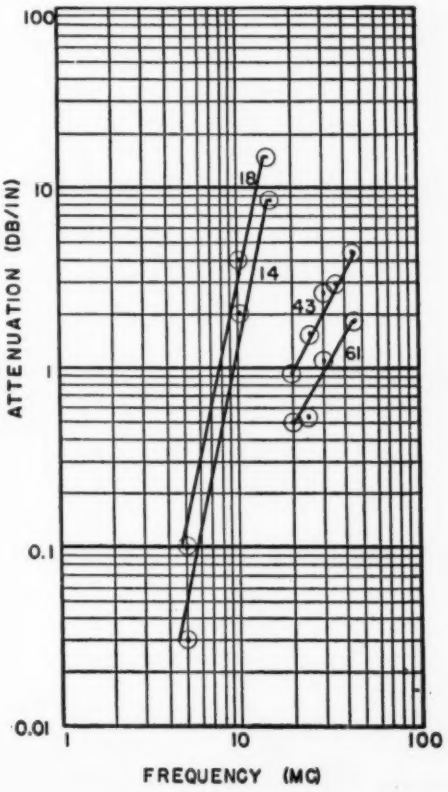


FIG. 19. Log-log attenuation-frequency plot for the 4150 steels.

in alignment. The accuracy of the measurements in steel is appreciably higher than in quartz because of the higher acoustical impedence of steel. The attenuation in the highly attenuating steel 14 and 18 samples, Fig. 17, shows only vaguely the curvature and break point associated with beam spreading. This is due to the relatively large amount of scattering which masks the beam spreading effects.

A linear plot of attenuation coefficient as a function of frequency for each of the steel samples is shown in Fig. 18. Figure 19 is a logarithmic plot of the linear portions of the curves in Fig. 18 and shows roughly that in steels 14 and 18 the attenuation coefficient is proportional to the fourth power of the frequency and that in steels 43 and 61 proportional to the square of the frequency.

Photomicrographs were made of sections of each of the steel specimens. After the attenuation measurements were made the specimens were cut up and those surfaces lying in or near the sound beam path were etched and photographed. The metallurgical history of all of the samples was the same as far as the austenitic grain size is concerned. The microstructure is brought out in the ×1500 photomicrographs, Figs. 20, 21, 22,

FIG. 20. ×750 photomicrograph of steel 14.

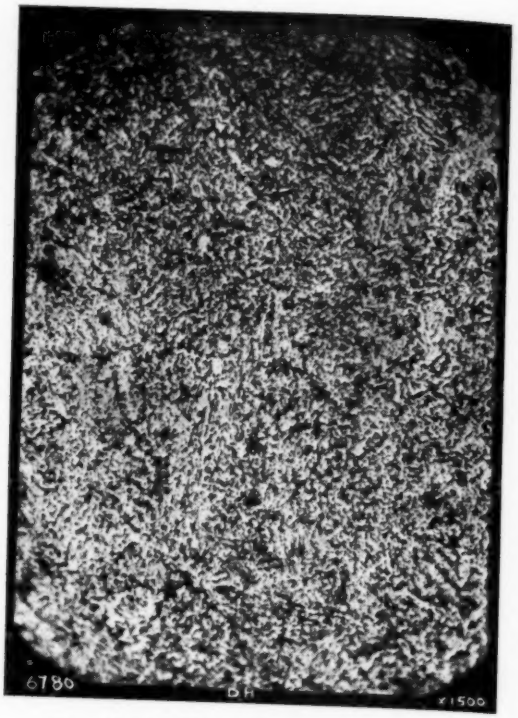


FIG. 22. ×750 photomicrograph of steel 43.

FIG. 21. ×750 photomicrograph of steel 18.

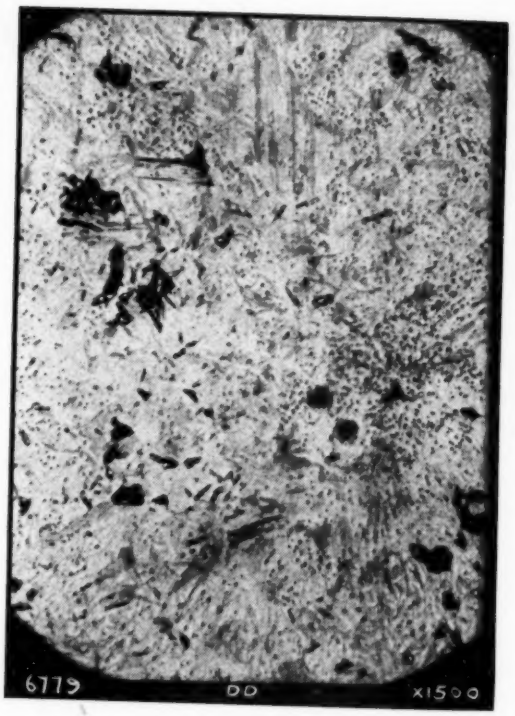


FIG. 23. ×750 photomicrograph of steel 61.

and 23. Metallurgical analysis of the microstructure is omitted here; it is apparent that the structural arrangement in samples 14 and 18 differ appreciably from sample 43 and in turn from sample 61. It should be pointed out that there are different types of etches and that the different types will show different features of the microstructure. It is sufficient for our purposes to say that samples 14 and 18 are very different in structure from samples 43 and 61.

While the number of specimens used in this work was

not large, some conclusions of physical and metallurgical interest can be drawn. Ultrasonic attenuation in the megacycle range is sensitive to differences in internal structure. Differences such as those caused in the steels by differences in heat treatment are very apparent as seen from the differences between the 14 and 18 steel on one hand and the 61 and 43 steel on the other hand. The 61 and 43 steel specimens were water quenched, and the 18 and 14 specimens were furnace cooled; steels 14, 18, and 61 were austenitized at the same temperature and have the same chemical composition, but $d\alpha/df$ differs by as much as a factor of 50. The frequency at which the attenuation becomes measurable by this technique is about 5 megacycles for the furnace cooled specimens and about 15 megacycles for the water quenched specimens. The sensitivity of the attenuation measurement to these differences is indicated by the difference detected between the two furnace cooled samples 14 and 18; at 10 megacycles there was a 100 percent increase in attenuation from the 14 specimen to the 18 specimen while the hardness index changed only 28 percent. It has been found that in some cases at least the measured differences are further enhanced when transverse waves are used instead of compressional waves.

The work done on polycrystalline aluminum and magnesium by Roth, Mason, and McSkimin, and Roney has shown how the scattering of ultrasonic waves by grain boundaries depends on frequency. In particular the scattering at low frequencies and small grain sizes appears to vary as f^4D^3. The attenuation *versus* frequency curves for steels 14 and 18 follow roughly a fourth-power law.

Sharp grain boundaries, such as appear in nonferrous metals, are not present in the photomicrographs of steel shown in Figs. 20 through 23, but regions of one structure do appear adjacent to regions of a different structure or perhaps adjacent to regions of the same structure differently oriented. Both gradual and sharp transitions from one oriented area or colony to another appear in the photomicrographs.

In the highly attenuating samples 14 and 18 there is in the photomicrographs indication of a high degree of anisotropy. This is in agreement with the general idea that a high degree of anisotropy of a material is associated with large scattering hence large attenuation.

Since steels 43 and 61 appear to be more homogeneous and isotropic than steels 14 and 18, the relatively low attenuation in the 43 and 61 samples is assumed to be related to their relative apparent homogeneity. The 43 and 61 samples have entirely different structures, and it will require further work to determine whether the difference in attenuation is caused by the degree of anisotropy or caused by the different kind of structure present.

Obviously a more extensive sampling program is necessary in order to make more than qualitative estimates of the causes of ultrasonic losses. The high sensitivity of the ultrasonic attenuation measurements suggests many uses in the physics of solids and in metallurgy as well as in quality control problems.

The methods described here in connection with the examination of steels are part of a larger program in this laboratory, both experimental and analytical, aimed at the study of the nature of scattering processes in single crystal and polycrystalline media under various electrical and mechanical conditions of the medium.

Support and encouragement for this and related work has been provided by a number of people at Watertown Arsenal Laboratory. In particular Mr. Carlton Hastings has provided continuous aid and advice. The metallurgical facilities and experience of Watertown Arsenal have been indispensable to this work. The contributions of Dr. H. Lester and Dr. R. Beeuwkes are also gratefully acknowledged as are those of all of the members of the Metals Research Laboratory at Brown University.

The Electric Field in Diodes and the Transit Time of Electrons as Functions of Current

PAUL L. COPELAND AND DELBERT N. EGGENBERGER
Illinois Institute of Technology and Armour and Company, Chicago, Illinois

(Received October 15, 1951)

Equations, applicable to all the most simple symmetries, are developed as approximate solutions of the related problems of potential distribution and transit time in diodes as functions of current. For the case of coaxial cylinders, calculations based on these equations have been made. Although these calculations fail to confirm Ivey's suggestion that the ratio of the field at the cathode with any degree of space-charge to that in the absence of space-charge is a universal function of the ratio of the actual current to the space-charge-limited current, they suggest that the function varies slowly as the geometry is changed. Thus small corrections to the function developed for parallel planes suffice to give a solution of high accuracy for any ratio of anode to cathode radius. Functions, applicable to electrons moving between coaxial cylinders, are tabulated to facilitate calculations of off-cathode field, off-anode field, transit time, and the potential distribution for any fraction of the space-charge-limited current.

WHEN electrons originating at an equipotential cathode travel in the field produced by their own space charge and the charges on the electrodes, the problems concerning potential distribution and transit time are interesting and important. For the most simple symmetries the limiting cases, in which space charge is negligible and in which the current is limited by space charge, have been solved.[1] In addition the potential distribution in the case of parallel planes has a satisfactory general solution which has been discussed recently by Ivey.[2] This solution has been extended by Barut[3] to take account of initial velocities common to all electrons. Recently also Page and Adams,[4] starting with a differential equation specifying the behavior of the potential distribution between coaxial cylinders, have developed solutions in terms of a number of infinite series, which taken together constitute a complete solution for the potential distribution between coaxial cylinders corresponding to any current per unit of axial length and any initial energy of the electrons injected into the space from the equipotential cathode. The variables chosen and the series obtained by Page and Adams do not make their expressions very convenient for the calculation of transit time, and the relation of transit time to the general problem has not been emphasized in the literature cited.

It is quite evident that the physical principles underlying the potential distribution for all of the simple symmetries are exactly the same, and it should be possible to formulate the solution in such a way that the equations developed could be applied to the entire range of problems arising for the various geometries and current densities. Thus the purposes of this paper are: first to show that for all of the simple symmetries, namely for parallel planes, coaxial cylinders, and concentric spheres, the solutions for the problems of potential distribution and transit time as functions of current density may be put upon the same formal basis by starting with Poisson's equation in the same form, and then from this formulation to derive additional relationships. The particular case in which electrons originate with negligible energy at an equipotential cathode will be emphasized, and detailed calculations will be confined to the practical situation of electrodes in the form of coaxial cylinders.

In the present analysis the dependence of space charge upon transit time is basic. The application of Gauss's law indicates that if the steady current between electrodes for any of these simple symmetries is multiplied by the transit time of the electrons, the product is the total instantaneous space charge. Similarly if the transit time in a definite part of the path is multiplied by the current, the result will be the corresponding part of the space charge. The transit time has definite limiting values which may be calculated for the case of negligible space-charge and for the case of space-charge-limited current.[5] The difference between the two limiting values amounts to about 50 percent. The transit time varies slowly as a function of space current, and the variation is especially slow for currents small in comparison with the space-charge-limited value. This situation in itself would make it possible to estimate the transit time for various currents, and it suggests that the transit time in the absence of space-charge is a basic function, the properties of which may be used for obtaining other desirable information such as the potential distribution with any fraction of the space-charge-limited current.

SOLUTION OF GENERAL PROBLEM

In all of the cases to be considered, the symmetries reduce the potential to dependence upon a single properly chosen coordinate. The boundary conditions take a particularly simple form if the distances in all cases are measured from a center of symmetry, and, because of this, the distance variable is called r. The law

[1] I. Langmuir and K. B. Blodgett, Phys. Rev. **22**, 347 (1923); **24**, 49 (1924); L. Page and N. I. Adams, Phys. Rev. **68**, 126 (1948); D. N. Eggenberger, M. S. Thesis, I. I. T. (1947).
[2] H. F. Ivey, Phys. Rev. **76**, 554 (1949).
[3] A. O. Barut, Phys. Rev. **81**, 274 (1951).
[4] L. Page and N. I. Adams, Phys. Rev. **76**, 381 (1949).
[5] P. L. Copeland and D. N. Eggenberger, J. Appl. Phys. **20**, 1148 (1949). Other symmetries are treated in Eggenberger's thesis.

of Gauss then gives

$$(dV/dr)A(r)=-4\pi\left[Q_c+\int_{r_c}^{r}\rho(r)A(r)dr\right], \quad (1)$$

where V is the electrostatic potential, Q_c is the total charge on the cathode, and $\rho(r)$ is the charge density in space. A is the area of a surface, defined by taking a constant value of r, through which all the electrostatic flux terminating on the cathode charge Q_c passes. Throughout the equations of this paper the cgs electrostatic units are used for all quantities. In general A will be a function of r, and in all of the simple symmetries it is a readily determinable function. We now take a very convenient coordinate z such that

$$dz=dr/A(r). \quad (2)$$

We then have

$$dV/dz=-4\pi\left(Q_c+\int_{r_c}^{r}\rho A dr\right)$$

and

$$d^2V/dz^2=-4\pi\rho A(dr/dz)=-4\pi\rho A^2. \quad (3)$$

This is Poisson's equation in a form which is very convenient for the subsequent analysis. Integration of Eq. (3) between any selected limits z_1 and z_2 gives

$$(dV/dz)_2-(dV/dz)_1=-4\pi q, \quad (4)$$

where q is the total charge between the surfaces for which $z=z_1$ and $z=z_2$.

Since the potential in the absence of space-charge is a linear function of coordinates, which like z, may be shown to be solutions of Laplace's equation simplified by the existing symmetries, these coordinates have a special advantage in that the equations assume a very simple form. We call them harmonic coordinates. As the need for additional symbols to represent this harmonic coordinate arises, we use letters such as u', y, and w, the differentials of which are all defined by the same relationship to dr, namely by Eq. (2).

When a steady current flows between electrodes of fixed difference of potential, the charge between any two surfaces in the space separating the electrodes, and defined by $z=z_1$ and $z=z_2$, is obtained through multiplying the current by the time t spent by each electron in passing from z_1 to z_2. Thus

$$q=it=[(dV/dz)_2-(dV/dz)_1]/(-4\pi), \quad (5)$$

where i is the current and t is the portion of the transit time which the electron spends between z_1 and z_2. Both i and q are negative by convention. If we agree to let I and Q represent the numerical magnitudes of i and q, respectively, Eq. (5) may be written without the negative sign before the 4π, and subsequent equations will be considerably simplified.

For the case of negligible current, the potential and transit time are functions of z which can be determined without great difficulty. The electrostatic potential V_0 is given by

$$V_0=C+kz, \quad (6)$$

where the constants C and k are to be evaluated from the boundary conditions. The increment ΔW in kinetic energy which the electron acquires between the cathode and a position corresponding to the coordinate z is $\Delta W=ek(z-z_c)$, where e is the numerical value of the electron's charge and z_c is the value of the harmonic coordinate at the cathode surface. The total kinetic energy of the electron is obtained by adding this to the initial kinetic energy, $mv_0^2/2$, where m is the mass of the electron and v_0 is the initial speed. In the case of negligible space charge we thus obtain

$$W=ek(z-z_c)+mv_0^2/2=eU_0(z), \quad (7)$$

where U_0 measures the total kinetic energy in units of potential difference. If all electrons have the same initial speed as they leave the cathode and if the velocity is directed along r, the transit time may be found from

$$t_0(z)=\int_{r_c}^{r(z)} dr/(2eU_0/m)^{\frac{1}{2}}, \quad (8)$$

where r_c is the value of r at the cathode and $r(z)$ is in general a value of r between the cathode and anode. For the complete transit time $r(z)$ is set equal to the plate radius.

The work of Ivey[2] and Barut[3] suggests that the field at the cathode is a simple function of I, and, letting $\delta\equiv I/I_s$, where I_s is the space charge limited current, we write the power series

$$(dV/dz)_c=(dV/dz)_0(1-a\delta-b\delta^2-c\delta^3-\cdots) \quad (9)$$

to express this relationship in a particularly convenient form.

In the case of steady currents uniformly distributed over surfaces of symmetry, current density is given by ρv. In the simple symmetries considered here, ρ and v are functions of the harmonic coordinate z. Since both v and ρ are constant for a fixed value of z, we may write

$$\rho=J/v=I/A(2eU(z)/m)^{\frac{1}{2}}, \quad (10)$$

where J represents the current density. If the function $U(z)$ were known, the transit time could be obtained with the help of Eq. (5). A method which permits the determination of $U(z)$ to any desired accuracy may be obtained from the equations

$$(dV/dz)_z=(dV/dz)_c+\int_{z_c}^{z}(d^2V/dz^2)dz, \quad (11)$$

$$V(z)=\int_{z_c}^{z}(dV/dz)_z dz. \quad (12)$$

In the foregoing equations A, I, I_s, $U_0(z)$, $t_0(z)$, $(dV/dz)_0$, and δ should be regarded as knowns. Obvi-

ously not all of the equations are independent, but taken together they completely determine the transit time and the potential distribution. The particular case in which the energy of the electrons at the cathode is negligible will be treated in detail. In this case $U_0=V_0$, and the solution for the electrostatic potential may be put into a more explicit form through inserting into the combination of Eqs. (11) and (12) values from Eqs. (3), (9), and (10) to give

$$V(z)=\left(\frac{dV}{dz}\right)_0(1-a\delta-b\delta^2-\cdots)\int_{z_c}^{z}dz + \frac{4\pi I}{(2e)^{\frac{1}{2}}/m}\int_{z_c}^{z}dz\int_{z_c}^{z}\frac{Adu'}{V^{\frac{1}{2}}}. \quad (13)$$

Our procedure in the calculation is suggested by the fact that the solution is known for δ equal to zero. Small currents will not alter the potential distribution much, and the charge distribution in the first approximation is given by $dQ=Idt_0$ where dQ is the part of the total charge associated with the part of the transit time dt_0. Having obtained the distribution of charge in the first approximation, the variation of the potential due to this charge is given at once by Poisson's equation, and the corrected expression for the potential may be inserted under the integral sign for use in the calculations. A solution with any desired accuracy may be obtained by carrying enough terms in the expression for the electrostatic potential. For illustration in this case of negligible energy at the cathode we carry terms up to and including order δ^2, δI, and I^2. This results in the expression

$$V(z)\simeq\left(\frac{dV}{dz}\right)_0(1-a\delta-b\delta^2)\int_{z_c}^{z}dz+\frac{4\pi I}{[2e/m(dV/dz)_0]^{\frac{1}{2}}}\int_{z_c}^{z}dz \times\int_{z_c}^{z}\frac{Adu'}{\left\{\int_{u_c'}^{u'}dw(1-a\delta)+\frac{4\pi I}{[2e/m(dV/dz)_0]^{\frac{1}{2}}}\int_{u_c'}^{u'}dw\int_{w_c}^{w}\frac{Ady}{y^{\frac{1}{2}}}\right\}^{\frac{1}{2}}}=\left(\frac{dV}{dz}\right)_0\int_{z_c}^{z}dz-\delta a\left(\frac{dV}{dz}\right)_0\int_{z_c}^{z}dz$$

$$+\frac{4\pi I}{[2e/m(dV/dz)_0]^{\frac{1}{2}}}\int_{z_c}^{z}dz\int_{z_c}^{z}\frac{Adu'}{\left(\int_{u_c'}^{u'}dw\right)^{\frac{1}{2}}}-b\delta^2\left(\frac{dV}{dz}\right)_0\int_{z_c}^{z}dz+\frac{2\pi I\delta a}{[2e/m(dV/dz)_0]^{\frac{1}{2}}}\int_{z_c}^{z}dz\int_{z_c}^{z}\frac{Adu'}{\left(\int_{u_c'}^{u'}dw\right)^{\frac{1}{2}}}$$

$$-\frac{m}{e}[2\pi I]^2\int_{z_c}^{z}dz\int_{z_c}^{z}\frac{Adu'}{\{(dV/dz)_0(u'-u_c')\}^{\frac{1}{2}}}\int_{u_c'}^{u'}dy\int_{y_c}^{y}\frac{Adw}{[V_0(w)]^{\frac{1}{2}}}. \quad (14)$$

In the final equality the first term on the right hand side gives the potential distribution in the absence of space charge. If the current is very small, the transit time may be computed as in the space charge free case. The charge is then determined from the relation $Q=It_0$, and the contribution of this charge to the potential is indicated by the terms containing the first power of the current or of δ. Thus the term in the expression for the off-cathode gradient which is linear in the current depends simply on the transit time in the absence of space charge. The computed change in the potential distribution as compared with the space-charge-free case has a small influence on the transit time, which in turn reacts to give a variation in the potential distribution of order I^2. Thus the term in the off-cathode gradient depending on the second power of the current involves only a more complicated function of the transit time in the absence of space charge and on the coefficient "a" as determined in the first approximation. It is clear that this process might be extended to give any desired accuracy in the potential distribution, because the coefficient of a given power of δ in the expression for the off-cathode gradient depends only on the coefficients of the lower powers of δ and on a determinable function of the transit time in the absence of space-charge.

A point of particular importance is that the boundary conditions are to be satisfied for all values of δ, and therefore the sum of the terms containing the first power of either δ or I must vanish when the upper limit of the integrals is assigned its value at the anode. This condition gives an equation for determining "a." Similarly the sum of the terms of order δ^2 must be zero for integrations extended to the anode, and this provides a means for the determination of "b." Similarly any other coefficients, such as "c" and "d," might be evaluated if desired.

The required integrals may be evaluated. If A depends on the zero power of z, as for parallel plane geometry, all integrations can be performed in closed form. Of course the potential distribution for parallel planes is known[2] and the suggested procedure is unnecessary in this case. The direct procedure is to evaluate the coefficients a, b, c, and d from Eq. (17) of Ivey's paper. Before this paper appeared, however, the writers, being unaware of the

existence of the exact solution of the problem, had applied the procedure outlined. The first two coefficients, which were obtained by this method, are in agreement with those subsequently obtained from Ivey's Eq. (17). Table I contains values computed from the latter equation.

For the case of concentric spheres, the leading terms depend on integrations which can be completed in closed form. The case of coaxial cylinders is the most difficult. The integrals have to be evaluated numerically with the aid of series. Because of the great practical interest in this case, we have chosen to illustrate the method by calculations for coaxial cylinders. In this case it is customary to reduce the area considered to that associated with unit axial length. Hence $du'=dr/2\pi r$ and $u'=(\ln r/r_c)/2\pi$. It is convenient to eliminate the numerical factor by using the variable u which is customary in this work. Thus $u\equiv 2\pi u'=\ln(r/r_c)$. Since the logarithm of unity is zero, $u_c=0$.

The Transit Time between Coaxial Cylinders without Space-Charge

The basic function in this analysis is the transit time of electrons between coaxial cylinders in the absence of space-charge. In the case of coaxial cylinders with internal cathode the potential function is

$$V=V_p(\ln r/r_c)/\ln r_p/r_c, \tag{15}$$

and the expression for the transit time of electrons leaving the cathode with negligible velocity is

$$t_0(x)=\int_{r_c}^{r} dr/v=r_c(mu_p/2eV_p)^{\frac{1}{2}}\int_1^x dx/(\ln x)^{\frac{1}{2}}, \tag{16}$$

where $x=r/r_c$. The problem presented here is that the integral appearing in Eq. (16) cannot be evaluated in finite terms. For small values of the ratio x, the variation of the logarithmic factor is most important, and this suggests successive integrations by parts in which the differential is chosen to be the logarithmic factor multiplied by the differential of the logarithm itself. In this way we obtain

$$T(x)=\int_1^x dx/(\ln x)^{\frac{1}{2}}=2\times(\ln x)^{\frac{1}{2}}\times[1-2(\ln x)/3+4(\ln x)^2/15-\cdots+\frac{(-2)^{n-1}(\ln x)^{n-1}}{1\cdot 3\cdot 5\cdots(2n+1)}+\cdots. \tag{17}$$

The series converges for all finite values of x, but except for low and intermediate values of x it is not very convenient because the calculations are very tedious. Well-known numerical methods for evaluating the contributions to the integral at the largest values of the variable x are more convenient, and they may be made sufficiently accurate for all practical purposes.

TABLE I.

Order of term	Symbol for coefficient	Value of coefficient
1	a	0.59259
2	b	0.109739
3	c	0.052025
4	d	0.031793
5	e	0.025195

If the cathode is the external cylinder

$$t_0=\left[\frac{m\ln r_c/r_p}{2eV_p}\right]^{\frac{1}{2}}\int_{r_c}^{r_p}\frac{(-dr)}{[\ln r_c/r]^{\frac{1}{2}}}.$$

Let

$$n=1/x$$

so that

$$dn/n=-dx/x \quad \text{and} \quad (-dr)=(r^2/r_c)dn.$$

Then

$$t_0=r_c[m(-u_p)/2V_pe]^{\frac{1}{2}}\int_1^n dn/n^2(\ln n)^{\frac{1}{2}}$$
$$=r_c[m(-u_p)/2V_pe]^{\frac{1}{2}}[(2/n)(\ln n)^{\frac{1}{2}}+(4/3n)(\ln n)^{\frac{1}{2}}+\cdots].$$

If in Eq. (17), x is replaced by $1/n$ and the square root of minus one occurring in the coefficient of the series is ignored, the result is identical with that just obtained. This rule will be used for the evaluation of the integrals by the series of Eq. (17). Values of this integral, which are fundamental to the calculation of transit time in the absence of space-charge, are shown in Table II.

Although the transit time in the absence of space-charge is the basic function, it enters the computation of potential as a result of further mathematical operations performed upon it. The function upon which the term containing the first power of the current depends is of the form

$$\int_1^x (dx/x)T(x).$$

The integrand is seen to be an infinite series. Each term of this series may be integrated by parts through repetition of the methods applied for the first integration. Thus we obtain

$$S(x)\equiv\int_1^x (dx/x)T(x)$$
$$=(4/3)xu^{\frac{1}{2}}[1-(4u/5)+(12/35)u^2-(32u^3/315)+(80u^4/3465)-\cdots]. \tag{18}$$

This function also is shown in Table II.

TABLE II.

$x=r/r_c$	$T(x)=\int_1^x \frac{dx}{u^{\frac{1}{2}}}$	$S(x)=\int_1^x \frac{dx}{x}T(x)$	$R(x)=\int_1^x \frac{dx}{u^{\frac{1}{2}}}S(x)$	$Q(x)=\int_1^x \frac{dx}{x}R(x)$	$u=\ln x$
0.00100	1.7720	11.355	−1.1386	6.88	−6.9078
0.0100	1.7682	7.280	−1.133	4.26	−4.6052
0.100	1.7159	3.245	−1.0576	1.705	−2.3026
0.200	1.6434	2.077	−0.9599	1.002	−1.6094
0.333	1.5274	1.2637	−0.8187	0.5441	−1.0986
0.500	1.3488	0.6768	−0.6292	0.2476	−0.69315
0.700	1.0664	0.2652	−0.3872	0.07393	−0.3567
0.800	0.8789	0.13456	−0.2611	0.03042	−0.2231
0.909	0.5983	0.0385	−0.1201	0.005832	−0.0953
0.990	0.1988	0.001321	−0.01332	0.0000658	−0.00995
1.00	0	0	0	0	0
1.01	0.2002	0.001326	0.01334	0.0000662	+0.00995
1.03	0.3473	0.006816	0.04012	0.0005895	0.02956
1.10	0.6376	0.04000	0.1346	0.006294	0.09531
1.25	1.0200	0.14713	0.3413	0.03637	0.2231
1.50	1.4688	0.3748	0.69736	0.1300	0.4055
2.00	2.1450	0.8943	1.4483	0.4322	0.69315
3.00	3.2003	1.9716	3.0820	1.323	1.09861
5.00	4.913	4.0204	6.774	3.737	1.6094
10.0	8.466	8.552	17.88	11.70	2.3026
20.0	14.58	16.33	46.03	32.33	2.9957
30.0	20.16	23.32	80.42	57.29	3.4012
50.0	30.59	36.08	164.5	117.2	3.9120
100.0	54.73	64.81	446.8	312.4	4.6052
200.0	99.51	116.66	1258.	854.0	5.2983
300.0	142.1	165.1	2345.	1561.	5.7038
500.0	223.9	257.0	5225.	3394.	6.2146
1000.0	418.6	472.5	15950.	10040.	6.9078

The final functions to be evaluated here are

$$R(x)\equiv\int_1^x (dx/u^{\frac{1}{2}})S(x)=(2/3)[x^2-1]-(8/3\cdot5)u^2$$
$$+x^2[(160u^3)/(3^2\cdot5\cdot7)-(272u^4)/(3\cdot5\cdot7\cdot9)$$
$$+(6304u^5)/(3\cdot5^2\cdot7\cdot9\cdot11)-\cdots] \quad (19)$$

and

$$Q(x)\equiv\int_1^x (dx/x)R(x)=\tfrac{1}{3}[x^2-1]-(2/3)u$$
$$-x^2[0.177777u^3-0.2158730u^4$$
$$+0.1439153u^5-0.0681097u^6+0.025502u^7$$
$$-0.007929u^8+0.002096u^9-0.000486u^{10}$$
$$+0.000100u^{11}-\cdots]. \quad (20)$$

Values of these functions also are listed in Table II.

Analysis of Results

It is of interest to determine "a" and "b" as functions of the ratio of the radii of the electrode cylinders. The space-charge-limited current per unit of axial length is

$$J'=\frac{2}{9}(2e/m)^{\frac{1}{2}}V_p^{\frac{1}{2}}/r_p\beta^2$$

where β^2 is a function which has been tabulated by Langmuir and Blodgett.[1] If in Eq. (14), I is interpreted as current per unit of axial length, $A=2\pi r$ and $I=\delta 2/9(2e/m)^{\frac{1}{2}}V_p^{\frac{1}{2}}/r_p\beta^2$. This value is to be used in Eq. (14). For a given pair of electrode cylinders with a given potential difference between them, Eq. (14) applied to obtain the potential of the anode should give the same result regardless of the value of δ. This condition requires that

$$a=[4u_p^{\frac{1}{2}}S(x_p)]/[9x_p\beta^2], \quad (21)$$

and

$$b=a^2/2-[8u_p^2Q(x_p)]/[81x_p^2\beta^4]. \quad (22)$$

For a given ratio of the radii of the electrodes "a" and "b" are constants, and in the application to a given geometry, they will be treated as constants. When from Eqs. (21) and (22) calculations are made for electrodes having various ratios of the radii, it is found that "a" and "b" are slowly varying functions of this ratio. The results of such calculations are shown in Table III.

It would be possible to carry terms involving the third and higher powers of δ in Eq. (14), and this procedure could be used to determine the behavior of "c." The labor involved in these calculations, however, increases quite rapidly as higher powers of δ are carried in them. At this point we can decide upon an approximation which is adequate for most practical purposes and which will involve very little additional labor. The proportionate variation in "b" is not greatly different from that in "a," but the former is only about a fifth as large as the latter. Therefore the importance of the variation of "b" in the determination of the field at the cathode is very much less than that of "a." Because a reverse

variation in the coefficients of the higher powers of δ is necessary to satisfy the known conditions imposed by the limiting cases, it is expected that there will be several coefficients in which the variation is extremely small. We decide, therefore, to neglect the variation in "c" and "d." The higher powers are required because the slope of the plot* of ϵ_c/ϵ_0 against δ becomes negatively infinite at $\delta=1$. It is not possible for a finite number of terms in the series for the ratio of the fields to represent the details of this behavior, but the effects remaining after the fourth-power term has been set down can be approximated by a single term of the seventh power in δ. Thus we propose as an adequate approximation

$$\epsilon_c=\epsilon_0(1-a\delta-b\delta^2-c\delta^3-d\delta^4-f_r\delta^7). \qquad (23)$$

In this equation a, b, c, and d should be regarded as knowns. It is also known that for $\delta=1$ the off-cathode gradient falls to zero. Thus for any ratio of the radii we may determine f_r from

$$f_r=1-a-b-c-d.$$

The values of the coefficient f_r listed in Table II were determined from this equation.

The Field at the Anode

In the parallel plane case Ivey[2] has shown that the ratio of the field at the anode with space charge to that in the absence of space charge is $\epsilon_p/\epsilon_0=[(\epsilon_c/\epsilon_0)^2+16\delta/9]^{\frac{1}{2}}$ and our own expansion of this in powers of δ gives

$$\epsilon_p/\epsilon_0=1+0.29630\delta+0.02195\delta^2+0.0065\delta^3+\cdots.$$

It will be useful to have a similar expression for the field at the anode in the cylindrical case.

Differentiation of the expression for the space-charge-limited current per unit of axial length gives

$$dV/du=\tfrac{2}{3}V_p[1+2(d\beta/du)/\beta]=\tfrac{2}{3}\epsilon_0 u_p[1+2(d\beta/du)/\beta].$$

The well-known behavior of β shows that for large positive values of u, the second term in the brackets becomes negligible, and for large values of u the ratio of the field at the anode with space-charge limitation to the field with negligible space-charge approaches the ratio $2u/3$. Term by term operations with the series for β results in the following expression for the ratio of the

TABLE III.

x	a	b	f_r	$(dV/du)_s/(dV/du)_0$	a'	b'	f_r'	u
0.00200	0.5907	0.1101	0.2154	1.0590	0.0513	0.0048	0.0029	−6.2146
0.01000	0.5909	0.1100	0.2152	1.0816	0.0702	0.0061	0.0053	−4.6052
0.1000	0.5920	0.1099	0.2143	1.1463	0.1287	0.0105	0.0071	−2.3026
0.2000	0.5923	0.1098	0.2141	1.1878	0.1621	0.0125	0.0092	−1.6094
0.3333	0.5924	0.1097	0.2141	1.2188	0.1943	0.0147	0.0096	−1.0986
0.500	0.5924	0.1097	0.2141	1.2545	0.2261	0.0169	0.0115	−0.69315
0.700	0.5925	0.1097	0.2140	1.2895	0.2573	0.0190	0.0132	−0.3567
0.800	0.5925	0.1097	0.2140	1.3050	0.2701	0.0200	0.0149	−0.2231
0.909	0.5926	0.1097	0.2139	1.320	0.2839	0.0211	0.0150	−0.0953
1.000	0.59259	0.109739	0.21385	1.33333	0.29629	0.02195	0.01509	0.0000
1.100	0.5926	0.1097	0.2139	1.3463	0.3075	0.0228	0.0160	+0.0953
1.25	0.5925	0.1097	0.2140	1.3657	0.3247	0.0241	0.0164	0.2231
1.50	0.5924	0.1097	0.2141	1.3928	0.3493	0.0259	0.0176	0.4055
2.00	0.5923	0.1097	0.2142	1.4421	0.3926	0.0293	0.0202	0.69315
3.00	0.5922	0.1097	0.2143	1.5228	0.4638	0.0352	0.0238	1.09861
5.00	0.5914	0.1097	0.2151	1.6493	0.5717	0.0444	0.0332	1.6094
10.00	0.5896	0.1097	0.2167	1.8619	0.7542	0.0612	0.0462	2.3026
20.0	0.5862	0.1094	0.2206	2.1149	0.9814	0.0840	0.0495	2.9957
30.0	0.5839	0.1092	0.2232	2.3200	1.133	0.1004	0.0865	3.4012
50.0	0.5801	0.1090	0.2271	2.5859	1.345	0.1238	0.117	3.9120
100.0	0.5733	0.1080	0.2349	2.9856	1.6564	0.1604	0.1688	4.6052
200.0	0.5650	0.1065	0.2447	3.420	1.9486	0.1980	0.274	5.2983
300.0	0.5593	0.1053	0.2516	3.709	2.187	0.2246	0.297	5.7038
500.0	0.5525	0.1041	0.2596	4.040	2.437	0.2556	0.347	6.2146
1000.0	0.5424	0.1014	0.2724	4.532	2.778	0.2975	0.457	6.9078

* *Note Added in Proof.*—In this paper epsilon is used to stand for the gradient in the harmonic coordinate. Thus the definitions are

$$\epsilon_0\equiv(dV/dZ)_0,$$
$$\epsilon_c\equiv(dV/dZ)_c.$$

Values of the integral involved in the computation of transit time in the absence of space charge can be obtained from standard tables through the substitution $x=e^{t^2}$. Then

$$\int_1^x \frac{dx}{(\ln x)^{\frac{1}{2}}}=2\int_0^{(\ln x)^{\frac{1}{2}}} e^{t^2}dt.$$

The values of this function with six-decimal-place accuracy for $.01\leq(\ln x)^{\frac{1}{2}}\leq 2.00$ were worked out by H. G. Dawson and are contained in Proc. London Math. Soc. **29**, 519 (1897). With four-place accuracy they are also given in Jahnke-Emde, *Tables of Functions* (Dover Publications, New York, 1943), p. 32.

off-anode gradients.

$$[(dV/du)_s/(dV/du)_0]_p = \tfrac{2}{3}[2+2.000000U+4.6666667U^2+6.54547U^3 +3.8067U^4-65.007U^5+46.807U^6 +177.8U^7+241.6U^8+85U^9\cdots \quad (23)$$

where $U=u/10$. When U is negative and not sufficiently small in magnitude to make the series converge rapidly, $d\beta/du$ may be evaluated from the assymptotic form given by Langmuir and Compton,[6] namely,

$$\beta^2 \simeq 4.6712(r_c/r)[\log_{10} r_c/1.4142r]^{\frac{1}{3}},$$

which gives

$$\left(\frac{dV}{du}\right)_s \Big/ \left(\frac{dV}{du}\right)_0 \simeq \frac{\ln r_c/r}{\ln .707 r_c/r}. \quad (23')$$

Values of this function are shown in Table III.

In general the ratio of the off-anode gradient with space-charge to that without space-charge is given by

$$\left(\frac{dV}{du}\right)_p \Big/ \left(\frac{dV}{du}\right)_0 = 1+a'\delta+\delta^2 b'+\cdots, \quad (24)$$

where

$$a' = [4u_p^{\frac{1}{3}}T(x_p)]/[9\beta^2 x_p]-a, \quad (25)$$

and

$$b' = \frac{2u_p^{\frac{1}{3}}T(x_p)}{9\beta^2 x_p}a - \frac{8u_p^{3}R(x_p)}{81x_p^2\beta^4}-b. \quad (26)$$

These functions also are shown in Table III. Since when $\delta=1$, the ratio of the gradients is given by Eq. (23), it is clear that we can determine the sum of the remaining coefficients in Eq. (24) from

$$f_r' = \left(\frac{dV}{du}\right)_s \Big/ \left(\frac{dV}{du}\right)_0 - a' - b' - 1, \quad (27)$$

and this also is shown in Table III.

Additional Relationships

In all cases we have $Q=It$, and differentiation gives

$$dQ = t\,dI + I\,dt.$$

This relation may always be used to obtain small increments in the charge from small increments in the current. The second term on the right hand side is usually small compared with the first, and a reasonably accurate estimate of dt will often suffice to give a result of high accuracy.

The Taylor's series expansion for the charge in powers of $(I-I_1)$ where $Q_1=I_1t_1$ is

$$Q = I_1t_1+[t_1+I_1(dt/dI)_1](I-I_1)+\cdots +[n(d^{n-1}t/dI^{n-1})_1+I_1(d^nt/dI^n)_1](I-I_1)^n/n!+\cdots.$$

If $I_1=0$, $(dQ/dt)_0=t_0$, and $(d^nQ/dI^n)_0=n(d^{n-1}t/dI^{n-1})_0$. The Maclaurin's series then is

$$Q = t_0I+(dt/dI)_0I^2+(d^2t/dI^2)_0I^3/2+(d^3t/dI^3)_0I^4/6+\cdots.$$

These relationships may be used directly for calculation, or through comparison with the previous relationships, still others may be obtained.

Use of the Relationships in Calculations

In all cases where a given potential difference is established between cathode and anode, the potential distribution in the absence of space-charge is easily determined, and the space-charge-limited current may be computed. Hence at any current which is physically possible, the current ratio δ may be determined. Equation (14) and the tables may then be used to approximate the potential distribution. The off-cathode field is given by Eq. (23) and may be evaluated with the aid of the tables. The off-anode field is given by Eq. (25) and the tables. The transit time for any degree of space-charge is then most easily computed by a proper application of Eq. (5).

CONCLUSION

Equations for the potential distribution as a function of the current, applicable to all of the simple symmetries —parallel planes, coaxial cylinders, and concentric spheres—have been developed. In this way it is emphasized that the physical principles in all of these cases are exactly the same. In this development, the behavior of electron transit time at low currents has been a fundamental consideration. The methods developed in this paper together with the tabulated functions for the case of coaxial cylinders have a great practical advantage. Although the numerical analysis underlying the study as applied to coaxial cylinders is quite tedious, the basic functions $T(x)$, $S(x)$, $R(x)$, and $Q(x)$ are considered to be of sufficient importance to have justified this work, and the completed tabulations provide a very convenient basis for calculations. Two circumstances contribute to the ease of calculations on this basis. First, the field at the cathode is a very slowly changing function of the geometry. Hence the solution which has been worked out for the plane parallel case is applicable with only small corrections, and an accuracy sufficient for all practical purposes is readily obtained. Secondly, the field at the anode, while it changes quite appreciably with geometry, can be expressed for each geometry in terms of a rapidly converging Maclaurin's series in powers of the ratio of the current to the space-charge-limited current. Thus Maclaurin's series expansions for the cathode field function, although slowly converging, can be approximated readily by corrections to the known behavior in the parallel plane case. The anode field, although far from a universal function, has a Maclaurin's series in powers of the current ratio which converges so rapidly that rather accurate information can be obtained without excessive labor from calculations based upon the first few terms. The fact that the principles employed are so general is in itself an aid in the application of this analysis.

[6] Langmuir and Compton, Revs. Modern Phys. 2, 247 (1931).

Letters to the Editor

Orientation of Crystallites in Stretched Polyethylene

SAMUEL KRIMM
Department of Physics, University of Michigan, Ann Arbor, Michigan
(Received November 5, 1951)

THE x-ray diffraction patterns of oriented polyethylene indicate that when the polymer is initially stretched at room temperature, the preferred direction of the crystallite chain axes does not coincide with the direction of elongation but is inclined to it.[1-3] As the elongation is increased, the crystallite orientation changes, until at elongations of about 500 percent the chain axes practically coincide with the stretching direction.[2] From an analysis of the x-ray diffraction patterns, Brown[2] has concluded that at elongations of up to 200 percent the long chain axis of the crystallite is inclined at an angle of 64° to the stretching direction. The purpose of this note is to point out that his analysis is not consistent with an angle of inclination of 64°, but rather the complement of this, *viz.*, about 26°.

After mentioning that the angle of inclination between the stretching direction and the long chain axis of the crystallites is 64°, Brown says, "There is a unique direction in the crystal lattice which is also inclined to the long chain axis at an angle of 64°. This is the 011 direction. Examination of Fig. 7 reveals that this reflection has indeed moved over to the meridian from its normal position in the first layer line. As a consequence of this, the 001 direction, or long chain axis, is inclined to the stretching direction at an angle of 64°."[4] The inconsistency in the above statement arises from a failure to distinguish between a direction in the real lattice, defined in terms of unit cell translations, and one in the reciprocal lattice, defined by the normal to a set of planes. Thus, the angle between the 001 and 011 directions in the real lattice of polyethylene, *viz.*, 62° 48′, is not the same as that between the vectors from the origin to the points 001 and 011 in the reciprocal lattice, which is 27° 12′. Therefore the movement of the 011 reflection to the meridian implies a tilt of 27° 12′ with respect to the stretching direction.

[1] W. O. Baker, *Advancing Fronts in Chemistry* (Reinhold Publishing Corporation, New York, 1945), Vol. I.
[2] A Brown, J. Appl. Phys. **20**, 552 (1949).
[3] S. Krimm and A. V. Tobolsky, J. Polymer Sci. **7**, 57 (1951).
[4] See reference 2, p. 555.

Reply to Samuel Krimm's Note

ALEXANDER BROWN
Carbide and Carbon Chemicals Company, South Charleston, West Virginia
(Received November 14, 1951)

THE comments made by Dr. Krimm about my paper are quité correct. I erred in the interpretation of the x-ray diffraction pattern. The angle of inclination between stretching direction and long chain axis of the crystallites is 26° (27°) and not 64°.

Distribution of Activation Energies for Viscoelastic Deformation

W. JAMES LYONS
Chemical and Physical Research Laboratories, Firestone Tire and Rubber Company, Akron, Ohio
(Received December 3, 1951)

ORIGINALLY defined to characterize macroscopic deformational phenomena, the concept of "relaxation time," along with the more recent idea of its distribution, is still in quite general use in the physical theory of polymers, even though more or less detailed models at the molecular level have since been developed. Essentially a mathematical abstraction of large scale effects, relaxation time becomes inadequate as a basic quantity when it is adapted to the molecular picture. The dependence of relaxation time on an over-all activation energy has been suggested by a number of writers[1-4] in accordance with the relationship:

$$\tau=\tau_0 \exp(F^*_{\mathrm{mol}}/RT), \tag{1}$$

where τ = relaxation time, τ_0 = constant, F^*_{mol} = activation energy per mole, R = gas constant, and T = absolute temperature. Specific experimental evidence is cited by deBruyne[2] in support of Eq. (1).

For a number of years, also, the distribution of relaxation times has been recognized qualitatively as an implication of the distribution of activation energies, as is suggested by the comments of Huggins.[5] It is the purpose of the present note to call attention to a relationship which (differing slightly from Eq. (1)) refers to the elementary (or unit) deformation process, and to present the distribution of activation energies corresponding to one experimentally derived relaxation-time distribution.

In terms of the reaction-rate theory, the shear modulus is given by Tobolsky, Powell, and Eyring[6] as

$$G=(\lambda_1 2\pi^2 F^*)/(\lambda^2 A), \tag{2}$$

where λ = distance between adjacent equilibrium positions, λ_1 = lateral distance between shearing molecular chains or segments, A = effective area of the deformation unit, and F^* = activation energy per molecule. The coefficient of viscosity is given by

$$\eta=(\lambda_1 h/\lambda^2 A)\exp(F^*/kT), \tag{3}$$

where h = Planck's constant, and k = Boltzmann's constant. If, now, we consider the rates of elastic and viscous shear in the molecular process to be additive, as do Tobolsky *et al.*, there is obtained an equation of the form of the well-known relaxation equation of Maxwell, according to whom the relaxation time is defined $\tau=\eta/G$. Thus, from Eqs. (2) and (3),

$$\tau=(h/2\pi^2 F^*)\exp(F^*/kT). \tag{4}$$

For constant F^*, Eq. (4) gives the same dependence of relaxation time on absolute temperature as does Eq. (1), but it should be noted that Eq. (4) gives the relaxation time of the *elementary* deformation process, whereas Eq. (1) applies to the viscoelastic system as a whole. It is significant that in the derivation of Eq. (4), other molecular parameters cancel out, and τ is dependent only on the activation energy at constant temperature.

In an analysis of dynamic-property and stress-relaxation data on polymers, Tobolsky, Dunell, and Andrews[7] point out that the dynamic modulus E of the whole system as a function of the continuous distribution of the relaxation times of component Maxwell units is given by the relation

$$E=\int_0^\infty E(\tau)d\tau, \tag{5}$$

where $E(\tau)$ is a distribution function such that $E(\tau)d\tau$ expresses the contribution to the total E, from units having relaxation times between τ and $\tau+d\tau$. Tobolsky and co-workers cite experimental results which indicate that

$$\left.\begin{aligned} E(\tau)&=E_0/\tau \quad (\tau_l<\tau<\tau_m),\\ E(\tau)&=0 \quad (\tau<\tau_l,\ \tau>\tau_m),\end{aligned}\right\} \tag{6}$$

where E_0 = constant. We may suppose each Maxwell unit to correspond to an elementary molecular deformation process. Then, referring to Eq. (4), it can be readily deduced that

$$E(\tau)d\tau=E_0(1/kT-1/F^*)dF^*. \tag{7}$$

The distribution function for activation energies $\mathcal{E}(F^*)=E_0(1/kT-1/F^*)$, corresponding to $E(\tau)=E_0/\tau$, is plotted in Fig. 1, for $T=300°\mathrm{K}$, $\mathcal{E}$ being expressed in multiples of E_0. Because of the nature of the function given by Eq. (4), in general two limits on F^* correspond to each of the limits τ_l and τ_m. In the present case, however, two of the F^* limits are superfluous, for they are below the lower limit ($F^*=kT$) which must be imposed if $\mathcal{E}(F^*)$

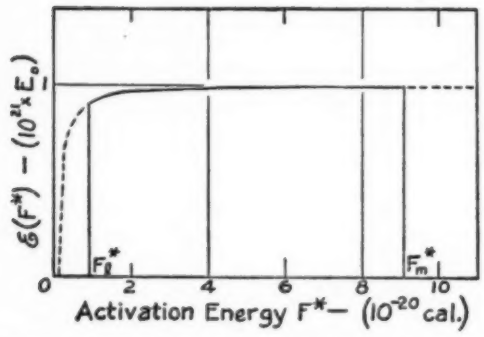


FIG. 1. Distribution function for activation energies, for $T=300°K$, based on the relaxation-time distribution given by Eq. (6).

is to have only the positive values for which there is ready physical interpretation. The two remaining limits F^*_l and F^*_m, corresponding to τ_l and τ_m, are shown in Fig. 1, being positioned quite generally. It is conceivable that F^*_l coincides with the limit $F^*=kT$, but this does not appear to be generally necessary.

The significance of $\mathcal{E}(F^*)$, as derived here, is analogous to that of $E(\tau)$. $\mathcal{E}(F^*)dF^*$ expresses the contribution to the dynamic modulus of the system from molecular deformation processes having activation energies lying in the interval F^* to F^*+dF^*.

[1] J. Frenkl, Z. Physik **35**, 652 (1926).
[2] N. A. deBruyne, Proc. Phys. Soc. (London) **53**, 251 (1941).
[3] Robert Simha, J. Phys. Chem. **47**, 348 (1943).
[4] T. Alfrey, *Mechanical Behavior of High Polymers* (Interscience Publishers, Inc., New York, 1948), p. 77.
[5] M. L. Huggins, J. Appl. Phys. **21**, 518 (1950).
[6] Tobolsky, Powell, and Eyring, edited by Burk and Grummitt, *Chemistry of Large Molecules* (Interscience Publishers, Inc., New York, 1943), p. 133.
[7] Tobolsky, Dunell, and Andrews, Textile Research J. **21**, 404 (1951).

Viscous Energy Dissipated During the Atomization of a Liquid

G. W. MONK
Camp Detrick, Frederick, Maryland
(Received November 28, 1951)

WHILE the detailed mechanism of generation of small droplets from bulk liquid is not completely understood, the difficulty in creating them at a high rate is readily apparent. When the liquid changes from the initial continuous state to the final dispersed state, its dimensions change by a factor of about 10^4. We shall assume for this discussion that the continuous state changes shape by flowing from the bulky configuration to a greatly extended thread or film which collapses under surface tension to droplets of various sizes, the mass average of which is at least as great as the thread or film thickness. Since the discrete state does not occur until after the change of shape, viscous energy must be dissipated in creating this extension and we shall show it to be appreciable in the rapid generation of small droplets.

The difficulty of obtaining the necessary rate of change of shape is illustrated by the fact that a one-micron thread would have to be ejected at about 10^8 cm/sec in order to disperse one cubic centimeter in a second. One way to estimate the minimum energy needed to effect the change of shape is to assume first that the liquid flows steadily into a transition region at low velocity and out at high velocity and small cross section. The viscous energy loss can then be computed by integrating the expression for the power loss per unit volume, W', at any point,[1] where

$$W'=\eta\{2(\partial u/\partial x)^2+2(\partial v/\partial y)^2+2(\partial w/\partial z)^2+(\partial u/\partial y+\partial v/\partial x)^2 +(\partial u/\partial z+\partial w/\partial x)^2+(\partial v/\partial z+\partial w/\partial y)^2\}, \quad (1)$$

in which η is the viscosity and u, v, w are the velocity components. If pure viscous flow occurs, a knowledge of the velocity distribution throughout the volume would allow a rather precise evaluation of the energy dissipated.

Since all the terms in Eq. (1) are positive, we shall attempt to estimate the magnitude of the first one, neglecting the others, in order to determine the minimum energy that must be dissipated in atomizing unit volume of liquid. If the liquid enters the large end of a conical transition region of diameter d_1 with a velocity v_1, suffers a constant rate of increase of velocity, and leaves as a thread of diameter d_2, the minimum energy dissipated per unit volume dispersed is of the order of W, given by

$$W=8\eta d_1^2R/3\pi d_2^4L, \quad (2)$$

where R is the rate of flow and L is the length of the transition region. The proportionality of W to R in Eq. (2) is in agreement with a more general rule, pointed out by J. B. Keller of New York University,[2] that in viscous flow the energy loss is proportional to the mass rate of flow. In order to give an idea of the magnitude of the energy, Eq. (1) has been used to compute the data of Table I.

TABLE I. Minimum viscous energy in calories required to atomize one cubic centimeter of liquid at a rate of one cm³/sec. The viscosity is assumed to be 0.01 poise, the length of the transition region and the diameter of its entrance to be 1 cm.

Thread diameter d_2, microns	Number of threads 1	10	100
1	2,000,000	200,000	20,000
5	3200	320	32
10	200	20	2.0
20	13	1.3	0.13

The resulting average droplet size should be at least as large as the thread diameter. Those values below the solid line in Table I probably represent conditions under which viscous energy losses are not a limiting factor in atomization. The existing two fluid atomizers operate under conditions such that this minimum viscous energy loss is not appreciable even though large quantities of energy are used in overcoming the resistance to flow of the propellant. It would therefore seem very difficult to generate a fine dispersion rapidly with the same inefficient method.

The single fluid hydraulic nozzle applies its available energy more efficiently to changing the shape of the liquid, but pressures of the order of 60,000 psi would be needed to supply 100 calories for each cm³ dispersed and much of this must be used in overcoming the resistance to flow caused by the stationary walls of the nozzle.

Although these approximations are very crude, they do seem to indicate the increasing importance of viscous energy losses as the droplet size is decreased. We are attempting to evaluate the magnitude more precisely in some particular cases.

[1] L. M. Milne-Thomson, *Theoretical Hydrodynamics* (Macmillan and Company, Ltd., London, England), second edition, p. 510.
[2] Private communication.

A Note on "Electron Microradiography of Electrodeposited Metals"

SHIGETO YAMAGUCHI
Scientific Research Institute, Bunkyo-ku, Tokyo, Japan
(Received December 3, 1951)

IT was reported by R. Weil and H. J. Read[1] that the electrodeposited thin film of nickel on copper and zinc served for study on electroplating by means of electron microscopy. It is known by electron diffraction investigation[2] that the crystal particles of metal, which are deposited electrolytically on the surface of a metal single crystal, are crystallographically oriented according to the orientation of the boundary faces of substrate. This phenomenon is quite similar to the oxidation taking place on metal; as a matter of fact, the thin oxide film formed on metal is composed of the oriented crystal particles. This oriented oxide film serves as a faithful replica of substrate in electron microscopy.[3] It is, therefore, reasonable to think that the electrode-

posited films prepared by Weil and Read express not only the initial state of plating, but also they are the replicas of the cleaned surfaces of copper and zinc substrata. The electrodeposited replica should be compared with the collodium replica prepared on the cleaned surface before electrodeposit. The discussion on the structure of replica film proposed in a previous letter[3] by the author might be applicable to the case of electrodeposited film.

[1] Rolf Weil and Harold J. Read, J. Appl. Phys. **21**, 1068 (1950); cf. G. M. Corney **22**, 682 (1951).
[2] G. P. Thomson and W. Cochrane, *Theory and Practice of Electron Diffraction* (Macmillan and Company, London, 1939), p. 170.
[3] S. Yamaguchi, J. Appl. Phys. **22**, 1295 (1951).

Radiation or Diffraction Patterns Close to Receiving Antennas

L. S. PALMER
The University College of Hull, England
(Received November 26, 1951)

IN the April issue of this journal C. L. Andrews described some interesting experiments with 8-cm waves in which a probe was used to measure the field intensity in the neighborhood of a half-wavelength rod when actuated by an incident plane electromagnetic wave from a distant transmitter. The field at any point was assumed to be the resultant of the incident wave and the wave reradiated from the rod. The field "pattern" round the rod was then obtained by considering the variation of the resultant field along any radial line. Andrews' experimental measurements were made, in particular, along the line joining the rod to the transmitter and along the line through the rod at right angles to this.

It is surprising to realize that Andrews' results with a single rod and 8-cm waves were, in fact, obtained in 1929 by Englund and Crawford[1] in America and at the same time by Palmer and Honeyball[2] in England, the former using a wavelength of 434 cm and the latter a wavelength of about 800 cm. In 1938 Palmer and Honeyball's work was extended by Palmer, Abson, and Barker[3] on a wavelength of 80 cm. In all these cases the measuring probe used by Andrews was replaced by a second tuned antenna. The results, as is to be expected, were practically identical as shown by the following figures photographically reproduced to the same scale of abscissas for convenience of comparison. Figure 1 of this letter compares Andrews' Fig. 2 with Figs. 5 (reversed) and 7 from reference 3 for the "end on" position ($\theta=0°$ and 180°), and Fig. 2 of this letter compares Andrews' Fig. 3 with Fig. 3 of reference 3 for the "broad-side on" position ($\theta=90°$).

Palmer and Honeyball's theory was based on the original dipole theory of Hertz; Andrews' theoretical equation (Eq. 6) is based on an expression for the radiation field given in Stratton's book,

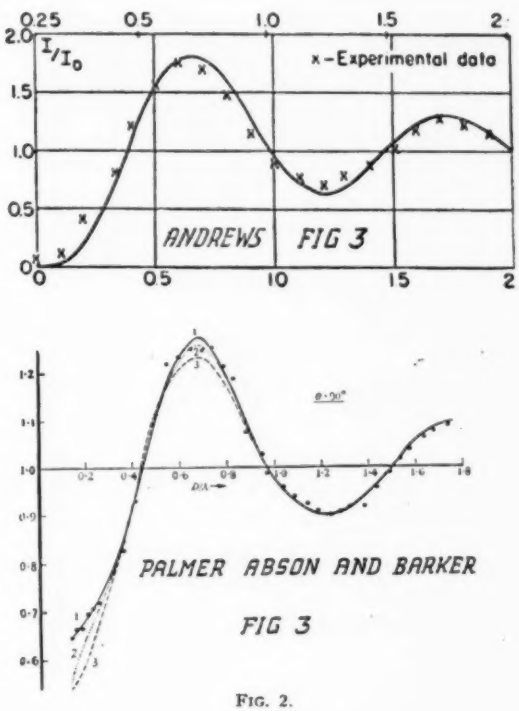


FIG. 2.

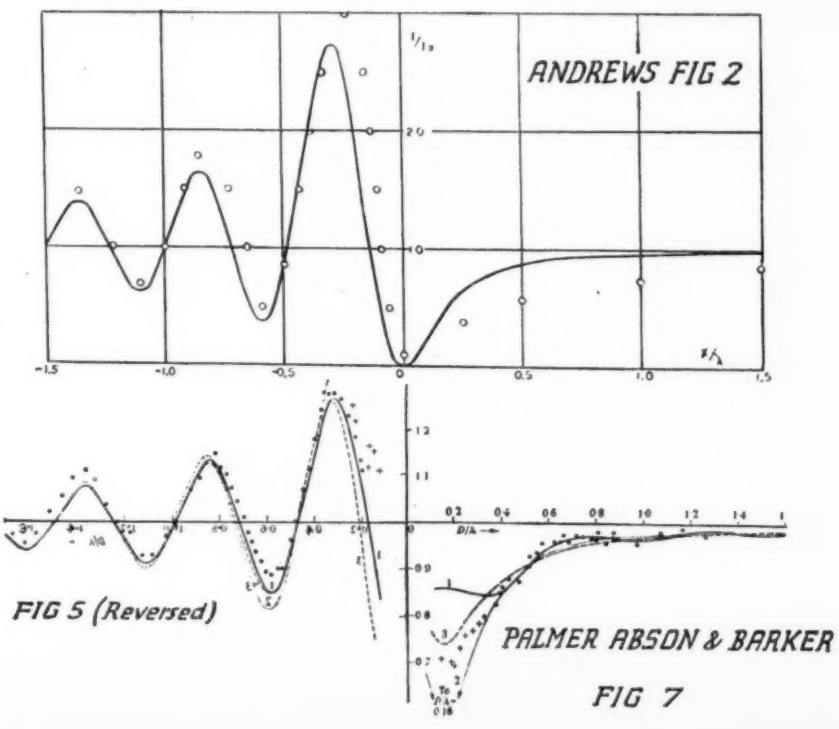


FIG. 1.

Electromagnetic Theory. They are, in fact, different treatments of the same fundamental electromagnetic problem.

By way of a further comparison, Palmer and Honeyball speak of the "reflecting" action of an antenna arising from initial absorption and subsequent reradiation of the incident wave, while Andrews refers to "the reradiated wave from the rod." Both these points of view are obviously the same; they yield the same results and both treat the problem as one of interference.

It is particularly interesting to realize that, after a lapse of twenty-two years during which the wavelengths used changed from 800 cm to 8 cm, there has been no fundamental change in the theoretical explanation of the results.

There is, however, a difference between the two points of view concerning the physical phenomena involved. Andrews refers to the "diffraction pattern" round the half-wave rod although he correctly explains the nature of the field as one produced by the interference of an incident wave from the transmitter with a reradiated wave from the rod. Does, in fact, a diffracting rod act as a reradiator of electromagnetic energy or does it act merely as an obstructive object absorbing energy from the incident radiation? It would perhaps be best to restrict the term "diffraction" to cases where opaque bodies absorb energy from an incident field with negligible reradiation. The problem of the nature of the residual field in the neighborhood of the diffracting body can then be treated in the usual way as a Fresnel or Fraunhofer diffraction problem without reference to the electrical properties of the body.

When, however, absorbed energy is reradiated, the interference between the reradiated energy and the incident energy may produce a resultant field quite different from the "residual" field arising from diffraction. Obviously, this effect will be greatest when no absorbed energy is degenerated to heat in the body. Thus the electrical properties of the body will be of paramount importance. Ideally, perfect conductors with maximum conduction current and perfect insulators with maximum displacement current will produce the maximum interference with the incident radiation in the neighborhood of the reradiating conductor or dielectric. There will obviously be no firm line of demarkation between these radiation fields and the diffraction fields of opaque nonradiators, because resistive conductors or conducting dielectrics will produce fields which will be combinations of the two.

It is clear, therefore, that radiators of small dimensions compared with the wavelength, even if they have imperfect electrical properties, will have negligible effect in absorbing incident radiation. Both Andrews and Palmer make this assumption. On the other hand, absorbing bodies of large dimensions compared with the wavelength will have negligible interfering reradiation but maximum obstructive absorption.

Thus the radiation problems of the wireless engineer would seem to be distinct from the optical problems of the physicist. It is not felt, therefore, as Andrews suggests, that a study of the radiation patterns produced by ultra-high frequency wireless waves will assist appreciably towards the understanding of optical diffraction problems. However, if the dimensions of the wireless aerial system could be enormously increased and its reradiating power similarly reduced, then ultra-high frequency diffraction fields would result but with little profit to the radio engineer. Alternatively, antenna radiation patterns might possibly help the physicist to understand those optical problems which can arise with ultra-microscopic absorbing and scattering obstacles. But the combined interference and diffraction fields that would be produced in these particular cases are an intermediate region in which very little experimentation has yet been undertaken.

[1] E. C. Englund and A. B. Crawford, Proc. Inst. Radio Engrs. **17**, 1277 (1929).
[2] L. S. Palmer and L. L. K. Honeyball, J. Inst. Elec. Engrs. **67**, 1045 (1929).
[3] Palmer, Abson, and Barker, J. Inst. Elec. Engrs. **83**, 424 (1938).

Concerning the Letter by Mr. P. E. Ohmart on "A Method of Producing Electric Current from Radioactivity"*

W. H. WALDO
Mound Laboratory, Monsanto Chemical Company, Miamisburg, Ohio
(Received January 31, 1952)

MR. OHMART'S professional affiliation at the time the work described in his letter was done was incorrectly given on page 1504. It should have read "Mound Laboratory, Monsanto Chemical Company, Miamisburg, Ohio." This work was performed under an AEC Contract Number AT-33-1-GEN-53. Mr. Ohmart's present address is the Ohmart Corporation, Cincinnati, Ohio.

* P. E. Ohmart, J. Appl. Phys. **22**, 1504–1505 (1951).

Books Reviewed

Ferromagnetism

By Richard M. Bozorth. Pp. 968+xvii. D. Van Nostrand Company, Inc., New York, 1951. Price $17.50.

For a number of years the only existing "bibles" of magnetism were either not up to date or not in English. Thus a need for new books in this field is great not only among research physicists who are interested in the fundamental aspects of theoretical and experimental magnetism but also among applied physicists, electrical engineers, metallurgists, etc. The long expected and recently published book by Bozorth was undoubtedly written with this widely diversified need in mind, and as a result it leans towards being an encyclopedia rather than a textbook. This is apparent even at a first glance through its 860 pages of actual text and some 100 pages of appendices, indexes, etc., which disclose an unusually high number of quantitative diagrams and references. Actually there are about one illustration and two references per page. In that respect the book is quite unique. No need to say that the appearance of a book of this type written by an outstanding researcher in the field of magnetism and a member of the technical staff of the Bell Telephone Laboratories, which for a number of years had a virtual monopoly in the field of magnetic research in this country, is an event of great importance.

In any book, and especially in a book which attempts to cover such a vast field, the choice of the boundaries and of manner of presentation is of vital significance. Bozorth's book consists essentially of two parts of about equal size. The first, of particular value to applied physicists, engineers, etc., is primarily a reference book and compendium of data on various magnetic alloys, and the subject matter is accordingly arranged in individual chapters on iron, on iron-nickel alloys, on iron-cobalt alloys, etc. Properties here described are mostly those of practical significance, and, especially for commercial alloys, they are very complete. A separate chapter is devoted to permanent magnet materials. Two introductory chapters deal with basic concepts of magnetism and with production problems such as melting, casting, role of impurties, etc.

The second half of the book is a survey of various magnetic phenomena and theories, and it also contains much experimental material. The accent is on a descriptive and intuitive manner of presentation with negligible amount of mathematics (for instance only 22 equations had to be numbered). "Derivations" are almost completely omitted and reference is given to original papers. The theory of magnetism proper, including a brief summary of the basic notions of paramagnetism, diamagnetism, and antiferromagnetism, is treated in about 50 pages. More than a quarter of this part of the book is a chapter on "Stress and Magnetostriction" which plays such a critical role for various practical applications and is of theoretical interest. The last chapter, 24 pages, deals with some of the more common and also some of the special methods of magnetic measurements.

As pointed out before, Bozorth's book is a huge and ambitious undertaking, and it attempts to cover a vast field of knowledge. Anyone who has tried to write a book of that sort, even on a much smaller scale, knows well what a thankless task it is to try to balance the choice of the material and method of approach so as to satisfy most of the readers. It is only by strict adherence to a certain pattern that a book acquires a character and remains within a reasonable size. Bozorth has accomplished this primarily by stressing the phenomenological aspects at the expense of theoretical analysis. This feature is no drawback (perhaps even an advantage) to a majority of those who are interested in applications and in engineering problems, and from that point of view the book is indeed very good. It is less satisfactory for those who would like to have a critical analysis and evaluation of current theories which, in the final instance, are the measure of our understanding of physical phenomena. The reviewer would have perhaps preferred a more extensive coverage of such interesting current topics as antiferromagnetism, ferrites, resonance, etc., and in particular more stress on the relation between magnetic and other phenomena in solids. These and similar wishes will be undoubtedly expressed by other readers, but it should be kept in mind that there is a limit to what can be included in a book (or, for that matter, in its review) even if it costs $17.50.

The book is very neat in appearance and nicely published. The subject index is unfortunately disappointing and uneven; for instance "resonance" is not listed, while names of various commercial alloys are. No need to say that in keeping with the best traditions of the Bell Telephone Laboratories the illustrations are very clear and well captioned. Altogether there is no doubt that the book will be of great value in the fields of magnetic research and magnetic engineering, both of which are recently receiving a renewed interest.

R. Smoluchowski
Carnegie Institute of Technology

Modern Interferometers

By C. Candler. Pp. 592. Hilger & Watts, Ltd., London N.N.W. 1, 1951.

This is a timely, comprehensive, lucidly expressed, and ably organized book on interferometry and its applications. Candler, whose books on spectroscopy are well known, has here carried out his most able exercise as an author, and this book can be recommended almost without qualification, not only to those who are interested in interferometers as tools, but to all who are interested in them as scientific instruments. An excellent balance is achieved between the theory of operation of interferometric instruments and their applications in the laboratory and shop.

After several chapters dealing with monochromatic radiation and the factors which limit monochromaticity, measurement of length in terms of wavelength is considered. Various methods of producing interference fringes are then discussed in detail, and the application of interferometry to measurements of all sorts is taken up. Measurements of displacements, of expansion, and of gauge thicknesses are considered in especially great detail. This discussion is followed by chapters on the Michelson interferometer and its applications, and on Twyman prism and lens interferometers. Gauge measuring interferometers are discussed in considerable detail, with the principal emphasis on British developments.

There are also chapters on the Fabry-Perot etalon and its uses. The Michelson stellar interferometer receives a chapter, and then the reflection and transmission echelons come in for careful discussion, as does the Lummer-Gehrke plate. Methods of adjustment of interferometers are gone into in such detail that the book will be found very useful by practical operators.

Chapter XV is concerned with the ruling of diffraction gratings and gives an excellent historical account, including a summary of work done in this field up to late 1947. Chapters are then devoted to the plane grating and the concave grating and to their uses and properties in various mountings.

While the reviewer is so impressed with the general excellence of this book that any adverse criticism seems captious, it is only fair to state that the author does not always say exactly what he means. In some cases this may lead to misapprehension on the part of the reader, as when the statement is made that "for this work (the measurement of length) the interferometer is superior to any other instrument, for it measures the differences in units which are invariable wavelengths of monochromatic light"; the barometer probably has more effect on a wavelength than on a foot rule. Similar carelessness extends in certain cases to the names of people, as when *Mack, Stehn, and Edlen* are on page 445 referred to as *Mach, Stern, and Edlen*. Some distortion in historical perspective might be induced by the reference to a grating "ruled by Rowland in 1913." Many staunch optical hearts will be perturbed by the statement on page 111 that "the Michelson interferometer may be fairly described as obsolete today; the

Twyman interferometer, however, is widely used by all who manufacture high quality optical instruments." Sufficient credit should be available for both Michelson and Twyman to make unnecessary specious arguments which suggest that the latter has displaced the former. The statement that "the Michelson interferometer is always made with the two arms at right angles, but this is not essential" seems unnecessarily severe, since most of those in the possession of the reviewer have their arms parallel, in one instrument these having been arranged so by Michelson in person.

In a field so large and growing, it would have been helpful to append initials to the authors' names given in the references, though in general these references are excellent. Although the above criticisms are representative of many others, they are, of course, minor, and would be easy to correct in the second edition, which is sure to be called for.

This reviewer was unable to find any type of interferometer that is not well described in the Candler book, and it should fill a very great need in a relatively new and rapidly expanding field, that of application of the most refined optical measurements to industrial uses.

GEORGE R. HARRISON
Massachusetts Institute of Technology

Elasticity

R. V. CHURCHILL, E. REISSNER, AND A. H. TAUB. Vol. III. Pp. 233+v, Figs. 10. McGraw-Hill Book Company, Inc., New York, 1951. Price $6.00.

"Elasticity" was the subject of the American Mathematical Society's Third Symposium on Applied Mathematics. This book contains the papers presented at that symposium. In "Approximate Methods of Solution of Two-Dimensional Problems in Anisotropic Elasticity," I. S. Sokolnikoff outlines several perturbational procedures and indicates that useful results can be obtained. "Beams under Concentrated Loading," by G. E. Hay, is a misleading title for a problem in plane strain of a cylinder whose section has confocal ellipses as boundaries and concentrated forces at the ends of the major axis of the outer boundary. From a practical point of view, the problem is left unsolved.

A masterly paper by E. Reissner, "On Axisymmetrical Deformations of Thin Shells of Revolution," is concerned with the nonlinear theory of finite deflections and is a major contribution to the theory of shells. F. E. Hildebrand, in a paper "On Asymptotic Integration in Shell Theory," considers small deflections of the general shell of resolution and presents asymptotic developments for the edge effects in explicit form suitable for direct application.

"Some Recent Applications of the Theory of Finite Elastic Deformation," by B. R. Seth, makes interesting reading because of its extraordinary claims and expressions of preference. In "Elastic Stability of the Facings of Sandwich Columns," H. W. March compares theories of surface wrinkling with each other and with experimental results. D. L. Holl describes a plan for attacking the problem of "Dynamic Loads in Thin Plates on Elastic Foundations."

By a systematic expansion of the stresses and displacements with respect to the thickness of the plate, K. O. Friedrichs succeeds in throwing new light and obtaining new results on the old problem of "Kirchhoff's Boundary Conditions and the Edge Effect for Elastic Plates." G. F. Carrier and F. S. Shaw present a remarkably brief and clear paper on "Some Problems in the Bending of Thin Plates" in which they describe an ingenious method for obtaining quick, approximate solutions for plates which are, or can be mapped into, regions bounded by a pair of intersecting straight lines and a closing curve. This is followed by an abstract of a paper on "Prestressing a Plane Circular Plate to Stiffen It against Buckling" by J. J. Stoker. A paper by C. L. Perry "On the Bending of Thin Elliptic Plates" describes the early stages of a solution in terms of Mathieu functions and another in terms of exponential functions.

Alexander Weinstein presents a very thorough and enlightening study of general theorems on inequalities and methods of obtaining upper and lower bounds in his paper on "New Methods for the Estimation of Torsional Rigidity." H. Poritzky, in a paper on "Stress Fields of Axially Symmetric Shafts in Torsion and Related Fields," gives solutions for an amazing variety of boundary shapes.

The remaining four papers, by W. Prager and P. S. Symonds, Sir Richard Southwell, N. Coburn, and E. H. Lee, are concerned with effects of plasticity.

On the whole, this book must be a source of gratification to those men who have been attempting to stimulate interest in elasticity and applied mathematics in general among mathematicians in the United States. Publication of a few of the papers could well have awaited completion of the authors' researches, but their premature appearance is not unusual in books of this type, as distinguished from the periodical literature.

R. D. MINDLIN
Columbia University

Theory and Application of Mathieu Functions

BY N. W. MCLACHLAN. Pp. 401+ix, Figs. 49. Oxford University Press, New York, 1950. Price $13.00.

The purpose of this book, as the author states in his preface, is to present the theory of Mathieu functions and to demonstrate its application to representative problems in physics and engineering science. The author further asserts that the book has been written for the technologist and is not addressed in any sense to the pure mathematician. The reviewer chooses to take this advice more as a declaration of purpose than as a literal fact, for it is unlikely that any technologist, however capable in his own field but naïve in the ways of mathematics, could more than begin to appreciate the wealth of detail which this book places at his disposal. On the other hand, the reviewer thinks it equally unlikely that any mathematician whose "pure" interests lay near the theory of Mathieu functions would find uninteresting the extensive development of properties and applications of this class of special functions contained in the present volume.

The book opens with a historical introduction tracing the evolution of the Mathieu function theory from the time of Mathieu's first memoir (1868) on the vibrations of an elliptical membrane. This section is followed by a sequence of thirteen chapters denoted as Part I—Theory of Mathieu Functions. The opening chapter, numbered II, develops the periodic solutions of the first kind for the canonical forms

$$(d^2y)/(dz^2)+(a\pm 2q\cos z)y=0$$

of Mathieu's equation, together with their orthogonality and normalization properties. Chapter III develops recurrence relations for the coefficients of solutions of integral order, studies convergence of series representations of solutions, and asymptotic behavior of coefficients as $q\to 0$, or $q\to\infty$, or $n\to\infty$. Chapter IV opens with a discussion of fundamental properties of the complete complement of solutions of the second-order linear differential equation with single-valued periodic coefficients. Solution functions of fractional order are introduced, and relations between the various forms of solutions are established. Chapter V follows with methods for obtaining numerical solutions of the Mathieu equation, illustrated by numerous examples. Chapter VI discusses Hill's equation and develops those cases in which the solutions involve Mathieu functions. Chapter VII develops the nonperiodic solutions which complement the periodic solutions of Chapter II to form a basis for a complete family of solutions. Chapter VIII relates the Mathieu equation to Bessel's equation and develops the expansion of Mathieu functions in series of Bessel functions.

In Chapter IX the two-dimensional wave equation is formulated in elliptical coordinates. The usual separation of variable technique leads to the canonical forms of the equations which came to bear Mathieu's name. The sequence of Chapters X–XIV develops numerous integral equations and integral relations satisfied by Mathieu functions and their modifications, together with asym-

ptotic formulas, studies on the location of zeros, and expansions in series of Bessel functions and products of Bessel functions.

Applications are explicitly discussed in Part II. In Chapters XV-XIX is given a succession of topics including amplitude distortion in loudspeakers, frequency modulation, vibrating strings with periodically varying end-point position, tension or mass, vibrations of elliptical membranes, plates, or water in elliptical-shaped lakes, viscous flow about an elliptic cylinder, and electrical and thermal diffusion in systems with elliptic shape properties. Chapter XVIII features the development of electromagnetic wave guide theory with particular application to guides with elliptical cross section. Chapter XIX studies the diffraction of sound and electromagnetic waves by elliptic cylinders.

The limiting forms of the various Mathieu functions as the fundamental ellipse degenerates into a circle or into a parabola are discussed in Appendix I. Characteristic numbers of a limited number of periodic solutions of integral order are tabulated in Appendix II. Appendix III classifies the various Mathieu functions according to the canonical forms of Mathieu's equation which they satisfy. An extensive bibliography is followed by a collection of new results obtained after the original manuscript had gone to press.

The reviewer considers the writing of this book an impressive achievement in working out and compiling that inevitable accumulation of detailed results which accompanies any class of special functions which become important in scientific and technological applications. The dilettante will not long browse in the pages of this book, but the serious worker whose interests lie in or near this field will find the book indispensable. That the author has written so well and the publisher produced such an attractive volume will make the use of the book all the more satisfying.

PAUL E. GUENTHER
Case Institute of Technology

Cloud Physics

BY D. W. PERRIE. Pp. 119, Figs. 7. John Wiley & Sons, Inc., New York, 1951. Price $4.50.

During the past few years considerable interest has been aroused in the field of cloud physics, particularly in connection with the possibilities of controlling the weather by means of artificial nucleation of clouds. This book apparently was written to present to its readers a review of some of the theories, old and new, concerning the physical processes involved in the formation and dissipation of clouds. The introduction of a chapter on induced precipitation, one of the first to be published, adds interest to the book, although many new advances have been made in this field since publication, and the reader should keep this in mind.

The first chapter is devoted to a detailed description of the various cloud types according to standard classifications. Apparently to help illustrate these descriptions, a section of 30 cloud photographs precedes the chapter.

Next, the composition of clouds and the physical processes by which they are formed are considered. Water droplet and ice crystal clouds are defined, followed by a discussion of the formation of clouds due to various lifting processes such as thermal, orographic, and frontal lifting and by other means, such as mixing, contact cooling, and the addition of water vapor.

Following this, there are chapters describing the various types of condensation and sublimation nuclei, natural precipitation processes, and methods of inducing precipitation artificially such as with the introduction of dry ice or silver iodide crystals into clouds. The list of references indicates that all sources of information on nuclei and induced precipitation had not been exhausted.

A chapter on the observation of clouds should prove valuable to those studying or observing clouds. Methods of determining cloud types are discussed, along with the methods used in determining cloud heights, velocities, etc. A section on cloud photography may be helpful for those interested in taking cloud pictures.

Practical application of a knowledge of cloud physics to forecasting various cloud formations, particularly in so far as aviation is concerned, is then brought out, and the last two chapters deal with the rare noctilucent and mother-of-pearl clouds and the optical and electrical phenomena associated with water droplet and ice crystal clouds.

The appendix contains useful charts of cloud characteristics. All pertinent information concerning the various cloud types is given in a handy reference table form. A glossary and extensive list of references complete an interesting book.

The author has attempted to cover a broad field in only 97 pages of text; consequently, most subjects are dealt with only briefly and in simple language without going into great technical detail. Nevertheless, *Cloud Physics* should prove useful to many readers, as a brief review of the subject to those who are established in meteorology, and as an introduction to the subject by those who are students or about to enter the profession.

RAYMOND E. FLACONER
General Electric Research Laboratory

Microwave Electronics

BY JOHN C. SLATER. Pp. 528+xiv. McGraw-Hill Book Company, Inc., New York, 1951. Price $7.50.

This book is particularly welcome because, until now, there has been no unified treatment on an advanced level of the electronic devices associated with microwave techniques. Included in this treatment are discussions of such devices as the klystron, magnetron, traveling-wave tube, and various particle accelerators. The author assumes a certain amount of mathematical sophistication on the part of the reader. For example, he solves the wave equation by use of expansions in orthogonal functions, but does not employ variational formulations or solutions in terms of Green's functions. In the discussion of particle accelerators, he derives the equations of motion from a Hamiltonian function.

Professor Slater has effectively divided the book into two parts: one half concerned with microwave circuitry and the other with microwave electronics. In the first section, he develops the solutions for the electromagnetic field in wave guides and proves their orthogonality properties. Then he considers the transmission line equations and from these develops the impedance concept and determines the power and energy relations. He presents a particularly complete analysis of resonant cavities, including in his treatment cavities with two or more outputs and problems connected with the coupling of cavities to wave guides. A logical extension is then made to the case of many-coupled cavities, i.e., the periodically loaded wave guide.

In the rest of the book Professor Slater deals directly with the interactions of charged particles and electromagnetic fields. After he describes the general properties of oscillators, amplifiers, and accelerators, he considers the specific example of the klystron. Here he develops the simple theory of bunching and discusses coupling of the bunched beam to the klystron cavity. In the section on the linear accelerator, the relativistic equations of motion are derived and the energy transfer expressions are developed. The problem of transverse stability of the electron beam is briefly discussed. For the traveling-wave tube, the propagation constant of the wave is found using a self-consistent field treatment.

Professor Slater discusses the magnetron at greater length. He considers first the electron motion in the static magnetron for both the linear and cylindrical case; then he discusses the space charge configuration quantitatively for the static magnetron and qualitatively for the oscillating magnetron. The many failures of the theory, including violation of the Hull cut-off condition, differing theoretical predications of the space-charge distribution in the static magnetron, etc., are not indicated. In the final chapter the author describes the synchro-cyclotron, betatron, and synchrotron and derives the relevant equations of motion. In addition, he discusses the stability of the particle orbits.

The conversational style and omission of many mathematical details make this book a very readable one. To pay for the latter advantage, the thorough reader may sometimes find it necessary

to spend considerable time between successive equations supplying these details on his own.

F. KARL WILLENBROCK
Harvard University

Physical Properties of Glass

BY J. E. STANWORTH. Pp. 224+vii, Figs. 107. Oxford University Press, New York, 1950. Price $4.25.

It is refreshing to find a book of this sort in a field where so much of the literature is descriptive. The author discusses structure, fracture, mechanical strength, diffusion of ions into glasses, electrical properties, surface characteristics, viscosity and surface tension, annealing, and devitrification. In all cases, the basic theoretical work in the subject and allied fields is critically considered. For example, the Griffith flaw theory is discussed in connection with mechanical strength. The volume is too small for an exhaustive examination of all the topics included, but the book should make an excellent starting point for a technologist who desires to learn. In this connection, numerous references to original literature are given. Typography and binding are of the usual high quality of this series of monographs.

F. B. HODGDON
American Lava Corporation

Structural Theory

BY H. SUTHERLAND AND H. L. BOWMAN. Pp. 394+xiv, Figs. 11.10. John Wiley & Sons, Inc., New York, Fourth Edition Revised, 1950. Price $5.00.

The fourth edition of this widely used book covers the essentials of the theory of statically determinate and indeterminate structures. The material has been partially rewritten, some topics have been extended, and a few have been added. The problems are unchanged. This book represents the approach used in most American engineering schools in the teaching of structural analysis, and it is to be lamented that a more advanced viewpoint with more emphasis on energy theorems is not more widely accepted.

M. G. SALVADORI
Columbia University

Laplace Transformation

BY WILLIAM T. THOMSON. Pp. 230+ix. Prentice-Hall, Inc., New York, 1950. Price $3.75.

This book is primarily concerned with applications of Laplace transformation to problems in mechanical engineering; as such, it should be of great interest and importance to the students in this field. From the standpoint of a physicist, however, the presentation lacks clarity somewhat. In particular, there is a noticeable lack of mathematical rigor in the proofs of the theorems; for example, in general there is not a careful delineation of such concepts as equation, function, and integral (cf. pp. 14, 24, and 38). In addition, there is some ambiguity in the use of concepts of analysis including reference to the Dirac function (cf. pp. 6, 25, 93, 95, and 167).

R. J. SEEGER
Naval Ordnance Laboratory

Response of Physical Systems

BY JOHN D. TRIMMER. Pp. 268+ix, Figs. 11.7. John Wiley & Sons, Inc., New York, 1950. Price $5.00.

In this book Professor Trimmer has presented a lucid generalized treatment of dynamic response from the differential equation point of view. The book begins with a philosophical introduction and a chapter devoted to definition of the terms employed. Chapters 3 and 4 treat first- and second-order linear systems in considerable detail. Sinusoidal forcing is discussed, and the inverse problem of finding the system parameters from the response to a known forcing is considered. A number of examples of these systems including a simplified version of a nuclear chain reactor are analyzed.

Chapter 5 discusses sinusoidal forcing of linear systems in general and introduces the ideas of transfer impedance, distortion, and Fourier analysis.

In Chapter 6 an approach to the study of higher order linear systems is presented. Third-order systems are covered in some detail. The problem of stability of higher order systems is considered. Several examples are treated including the nuclear reactor having a number of delayed neutron groups.

Chapter 7 is a discussion of measuring instruments as physical systems. Various characteristics of instruments are considered including accuracy and error which are treated in some detail.

Feedback systems are covered briefly in Chapter 8. The remaining chapters deal with parametric forcing, distributed systems, and nonlinear systems. The Laplace transform method, while not employed in the text, is outlined in an appendix.

The list of problems included as an appendix seems rather brief by comparison with the scope of material covered. The extension of this list together with a categorical arrangement of the problems would no doubt be helpful to those interested in using this book as a text.

The emphasis in this book is placed on the approach to the problem and the interpretation of the results. Particular attention is paid throughout to careful definition and use of terms. A unified system of symbols is used to emphasize the mathematical similarity of a wide variety of physical systems. This volume should prove particularly valuable to scientists and engineers who desire a comprehensive yet compact introduction to the growing field of "system response."

ROBERT I. STROUGH
Pratt & Whitney Aircraft

Ultrasonics

BY P. VIGOUREUX. Pp. 163, Figs. 74. John Wiley & Sons, Inc., New York, 1951. Price $4.00.

The discovery in gases of dispersion of sound velocity and anomalous bands of absorption at ultrasonic frequencies, with subsequent explanation in terms of intramolecular processes, initiated in recent years a wide interest in the use of sound as a means of studying the physical properties of matter. The author, who has had a broad experience in acoustic research, has chosen to confine his attention to this specialized, but highly interesting, phase of ultrasonics, the measurement of absorption and velocity. There is a need for a review covering work in this field during the past ten years and the author has made a welcome contribution.

Chapter III is an excellent, although brief, account of simple theory to describe attenuation and dispersion due to viscosity thermal conduction and intramolecular processes. Scattering and refraction are also mentioned. Molecular absorption is treated in a simple manner by giving part of the specific heat a time constant and deriving the frequency dependence of the specific heat from an electric circuit analogy. The frequency dependence of attenuation and velocity is then analyzed in detail and presented graphically.

The meat of the book is contained in the last two chapters, V and VI, where measurement work on liquids and gases is reviewed. Solids are not discussed. Calculations of the molecular heats of gases are presented to show that, at least for simple molecules, attenuation and dispersion can be predicted with fair accuracy. The work of Hall, describing absorption in liquids due to a structural relaxation, is presented in some detail. Here one regrets the lack of reference to the work of Eckart [Phys. Rev. **73**, 68 (1948)] and Liebermann [J. Acoust. Soc. Am. **20**, 868 (1948); Phys. Rev. **75**, 1915 (1949)] who have explained absorption in water by the usual hydrodynamic theory, taking into account a bulk viscosity, which they have measured, as well as the shear viscosity. In view of this work the statement (p. 126) "Thus at present only the monatomic liquids are left as instances where absorption is accounted for by viscosity···" is somewhat misleading.

Chapters II and IV (Chapter I is an introduction) present, in accordance with the author's purpose "to indicate general principles rather than enter into detailed description of apparatus or of experimental procedure," a sketchy but palatable description of experimental techniques and equipment. That classic friend of the acoustician, the piston radiator, is mentioned. Quartz crystal and magnetostriction transducers are adequately treated for the purpose, with some attention to motional impedance analysis. Among the methods by which velocity and attenuation can be measured, the interferometer, optical diffraction, and echo pulse methods are selected for discussion and comparison.

For details not included in this brief review and for material of a more general nature, there is appended an extensive bibliography covering work from January 1, 1939. A notable omission is reference to the interesting work on gases at low pressures where the wavelength can approach the free mean path. Greenspan has given references to this work [J. Acoust. Soc. Am. **22**, 568 (1950)].

BERNARD D. SIMMONS
National Bureau of Standards

Electromagnetic Waves

BY F. W. G. WHITE. Pp. 108+viii, Figs. 25. John Wiley & Sons, Inc., New York, 1950. Price $1.25.

This book is a recent, fourth revised edition of a monograph of one hundred small pages. It gives an efficient description of classical electromagnetic theory, starting with the basic experimental laws and Maxwell's equations, with emphasis on the Lorentz theory, and on refraction, absorption, diffraction, and propagation in an ionized region.

The choice of topics and the approach reflect the older rather than the newer trends. For example, no mention is made of the guiding of electromagnetic waves by mental and dielectric boundaries, nor of antennas, nor of propagation in the presence of a perfect or imperfectly conducting ground. Magnetic poles are used in the derivations. MKS units do not appear. The bibliographies at the end of each chapter include the classical texts and articles on the subject treated. No reference was noticed to articles written after 1934.

The quantity of material covered could hardly be greater for the size of the text; but, with the newer contributions more strongly influential, a monograph on the basic theory probably would have been differently written. On the other hand, the new material is now so commonly available in books with similar titles that it is worthwhile to have also this little text with its neat summary of the classical theory.

SIMON RAMO
Hughes Aircraft Company

Principles and Applications of Waveguide Transmission

BY GEORGE C. SOUTHWORTH. Pp. 689+xi, Figs. 12.5-5. D. Van Nostrand Company, Inc., New York, 1950. Price $9.50.

The uniconductor wave guide as a means of microwave transmission was rediscovered and first put into practical use in the thirties by the group at Bell Telephone Laboratories and Dr. Barrow of M.I.T. It is appropriate to have a book by one of the pioneers treating the subject and the related topics in a comprehensive manner. Although enough basic and elementary mathematical theory is given, the main emphasis is on the applications of the principles of wave guide to various microwave devices.

After a brief review of the history of wave guides, a detailed summary of the principles of networks and transmission lines is given. The next three chapters cover the wave guide theory. It is followed by descriptions and quantitative analyses of wave guide elements and components essential for a microwave transmission system. The second half of the book deals with the applications of wave guide principles to microwave antennas, electron tubes, and modulation and demodulation devices.

This book will serve as an excellent reference because of the extensive coverage of microwave devices presented mostly in descriptive form. Many of the devices were not well known to the public. The bibliography is fairly complete.

L. J. Chu
Massachusetts Institute of Technology

High-Speed Computing Devices

BY THE STAFF OF ENGINEERING RESEARCH ASSOCIATES, INC. Pp. 451+xiii, Figs. 15-5. McGraw-Hill Book Company, Inc., New York, 1950. Price $6.50.

This is a good book prepared by the staff of Engineering Research Associates under a contract with the Office of Naval Research. Material was collected by interviews and visits to the groups working actively in the computer field.

The book covers components and physical techniques of computing machines and is carefully done and technically sound. It does not attempt to treat the applications of computing machines and only briefly the subject of the complete machine. There is a chapter on numerical analysis. The book has an extensive bibliography divided by chapters.

Part I on Basic Elements of Machine Computation discusses counters, switches, and gates. Part II covers computing systems from desk calculators through large scale automatic digital computers. There is a chapter on analog computing; most of the book is on digital devices. Part III on Components and Methods discusses computing, storage, and analog-to-digital data conversion.

There are few books on mechanical and electronic computation, and this fills one of the vacancies by treating basic physical techniques.

JAY W. FORRESTER
Massachusetts Institute of Technology

Electromagnetic Fields: Theory and Application. Volume I: Mapping of Fields

BY ERNST WEBER. Pp. 590+xiv, Figs. 34.2. John Wiley & Sons, Inc., New York, 1950. Price $10.00.

Mapping of Fields is a text designed for graduate students of electrical engineering, but also intended for physics students. It treats the various theoretical methods for determining static electric, magnetic, and electric current fields. In the first two chapters the principles of electrostatic and magnetostatic fields are explained, while in the third chapter static electric current fields are described and brief mention is made of other physical fields. In the remainder of the book numerous problems are solved exactly or approximately by various methods, and many results are quoted from the literature. The methods of superposition, of images of inversion, of conformal mapping and separation of variables are described and applied. Graphical field plotting, numerical methods, and experimental mapping methods are also explained.

Although the book is quite long, much of it is cramped for space. For example, the solutions of many problems are merely sketched with references to other texts or to the literature, for complete solutions. Many topics are discussed so briefly that no idea of their use or value can be gained. It seems that the attempt to present the subject from scratch is incompatible with the attempt at completeness. On the one hand the elements of complex variable theory are developed beginning with the definition of a complex number. On the other hand many problems are followed by lists of texts and articles employing the same or similar methods, giving graphs or numerical results, etc. The latter information is invaluable to those already acquainted with the subject, and especially to research workers. But it seems out of place in a book for beginners.

In summary, the book is a good attempt at filling an obvious gap in the textbook literature.

JOSEPH B. KELLER
Institute for Mathematics and Mechanics
New York University

Linear Computations

By Paul S. Dwyer. Pp. 344+xi. John Wiley & Sons, Inc., New York, 1951. Price $6.50.

This book is concerned with the problem of solving systems of simultaneous linear algebraic equations and related problems on a desk computing machine. No mathematical knowledge is assumed beyond elementary algebra. First comes a treatment of the machine performance of various arithmetic operations and the first-order errors which result. This is followed by several chapters which contain descriptions of various methods of solving systems of linear algebraic equations, these methods having various degrees of practical utility; and the relations between the quantities computed in using the different methods is considered in detail. Corresponding methods for evaluating determinants are next described; also the evaluation of linear forms and the relationship between the solutions of systems of equations having certain similarities are considered. The next few chapters have to do with the calculation of the inverse matrix and the characteristic equation, after which there is one on the calculation of the first-order errors involved in carrying out the processes previously described. Finally certain results previously obtained pertaining to machine calculations are applied to various processes arising in statistics. The book contains many numerical examples.

Prescott D. Crout
Massachusetts Institute of Technology

Other Books Received

The Nature of Polyphase Induction Machines. Philip L. Alger. Pp. 397+xiii. John Wiley & Sons, Inc., New York, 1951. Price $7.50.

Advanced Fluid Dynamics and Fluid Machinery. R. C. Binder. Pp. 426. Prentice-Hall, Inc., New York, 1951. Price $6.00.

The Restless Universe. Max Born. Pp. 352, Figs. 120. Dover Publications, Inc., New York, 1951. Price $3.95.

Philosophical Problems of Mathematics. Bruno, Baron v. Freytag gen. Lorikghoff. Pp. 88. The Philosophical Library, Inc., New York, 1951. Price $2.75.

Statistical Methodology Reviews (1941–1950). Edited by Oscar Krisen Buros. Pp. not numbered. John Wiley & Sons, Inc., New York, 1951. Price $7.00.

The Earth's Magnetism. Sydney Chapman. Pp. 127+xi, Figs. 39. John Wiley & Sons, Inc., New York, 1951. Price $1.50.

Perturbation Methods in the Quantum Mechanics of *n*-Electron Systems. E. M. Corson. Pp. 308+xi. Hafner Publishing Company, Inc., New York, 1951. Price $11.00.

Musical Acoustics, Third Edition. Charles A. Culver. Pp. 215+xiv, Figs. 147. The Blakiston Company, Philadelphia, 1951. Price $4.25.

Ultraviolet Spectra of Aromatic Compounds. Robert A. Friedel and Milton Orchin. John Wiley & Sons, Inc., New York, 1951. Price $10.00.

Astrophysics. Edited by J. A. Hynek. Pp. 703. McGraw-Hill Book Company, Inc., New York, 1951. Price $12.00.

Some Aspects of Fluid Flow. The Institute of Physics. Pp. 292. Longmans, Green and Company, Inc., New York, 1951. Price $9.50.

High Frequency Transmission Lines. Willis Jackson. Pp. 152+viii, Figs. 46. John Wiley & Sons, Inc., New York, 1951. Price $1.75.

Astronomy of Stellar Energy and Decay. Martin Johnson. Pp. 216, Figs. 21. Dover Publications, Inc., New York, 1951. Price $3.50.

The Design of Switching Circuits. William Keister, Alistair E. Ritchie, and Seth H. Washburn. Pp. 556. D. Van Nostrand Company, Inc., New York, 1951. Price $6.00.

Fever Therapy. H. Worley Kendell. Pp. 101. Charles C. Thomas, Springfield, Illinois, 1951. Price $2.25.

An Introduction to the Theory of Control in Mechanical Engineering. R. H. MacMillan. Cambridge University Press, New York, 1951. Price $6.00.

Quantum Mechanics of Particles and Wave Fields. Arthur March. Pp. 292. John Wiley & Sons, Inc., New York, 1951. Price $5.50.

Fundamentals of Electronics. F. H. Mitchell. Pp. 243. Addison-Wesley Press Inc., Cambridge, Massachusetts, 1951. Price $4.50.

Electromagnetic Problems of Microwave Theory. H. Motz. Pp. 184+vii. John Wiley & Sons, Inc., New York, 1951. Price $2.00.

Simplified Mechanics and Strength of Materials. Harry Parker. Pp. 275. Figs. 140. John Wiley & Sons, Inc., New York, 1951. Price $4.00.

Isoperimetric Inequalities in Mathematical Physics. G. Pólyna and G. Szegö. Pp. 279. Princeton University Press, Princeton, New Jersey, 1951. Price $3.00.

Thermodynamics. A. W. Porter. Pp. 124+vii. John Wiley & Sons, Inc., New York, 1951. Price $1.50.

Time Bases, Second Edition. O. S. Puckle. Pp. 387, Figs. 257. John Wiley & Sons, Inc., New York, 1951. Price $5.00.

Introduction to the Study of Aircraft Vibration and Flutter. Robert H. Scanlan and Robert Rosenbaum. Pp. 428+x. The Macmillan Company, New York, 1951. Price $7.50.

Engineering Thermodynamics. Herman J. Stoever. Pp. 458, Figs. 140. John Wiley & Sons, Inc., New York, 1951. Price $5.75.

Integral Transforms in Mathematical Physics. C. J. Tranter. Pp. 118. John Wiley & Sons, Inc., New York, 1951. Price $1.50.

200 Miles Up. The Conquest of the Upper Air. J. Gordon Vaeth. Pp. 203, Figs. 62. The Ronald Press Company, New York, 1951. Price $4.50.

Primary Batteries. George Wood Vinal. Pp. 336+vii, Figs. 101. John Wiley & Sons, Inc., New York, 1950. Price $5.00.

Optical Crystallography, Second Edition Revised. Ernest E. Wahlstrom. Pp. 247+vii. John Wiley & Sons, Inc., New York, 1951. Price $4.50.

Radiologic Physics, Second Edition. Charles Weyl and S. Reid Warren, Jr. Pp. 491. Charles C. Thomas, Springfield, Illinois, 1951. Price $10.50.